

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>5</sup>:</b> <b>C07H 21/04, C12N 15/00, 1/20, C07K 1/00, C12P 21/06</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 99/47538</b> <b>(43) International Publication Date:</b> 23 September 1999 (23.09.99)
<b>(21) International Application Number:</b> PCT/US99/05068 <b>(22) International Filing Date:</b> 5 March 1999 (05.03.99)  <b>(30) Priority Data:</b> 60 078,563 19 March 1998 (19.03.98) US 60 086,505 22 May 1998 (22.05.98) US  <b>(71) Applicant (for all designated States except US):</b> HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US).  <b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> RUBEN, Steven, M. [US/US]; 18528 Heritage Hills Drive, Olney, MD 20832 (US); ROSEN, Craig, A. [US/US]; 22400 Rolling Hill Road, Laytonsville, MD 20882 (US); MOORE, Paul, A. [US/US]; 19005 Leatherbark Drive, Germantown, MD 20874 (US).  <b>(74) Agents:</b> BROOKES, A., Anders et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 20850 (US).	<b>(81) Designated States:</b> AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GR, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
<b>(54) Title:</b> CYTOKINE RECEPTOR COMMON GAMMA CHAIN LIKE  <b>(57) Abstract</b>  The present invention relates to a novel human protein called Cytokine Receptor Common Gamma Chain Like, and isolated polynucleotides encoding this protein. Also provided are vectors, host cells, antibodies, and recombinant methods for producing this human protein. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating disorders related to this novel human protein.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroun	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## Cytokine Receptor Common Gamma Chain Like

### *Field of the Invention*

The present invention relates to a novel human gene encoding a polypeptide which is a member of the Cytokine Receptor family. More specifically, the present invention relates to a polynucleotide encoding a novel human polypeptide named Cytokine Receptor Common Gamma Chain Like, or "CRCGCL." This invention also relates to CRCGCL polypeptides, as well as vectors, host cells, antibodies directed to CRCGCL polypeptides, and the recombinant methods for producing the same. Also provided are diagnostic methods for detecting disorders related to the immune system, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying agonists and antagonists of CRCGCL activity.

### *Background of the Invention*

Biochemical and physiological effects often result from the binding of a cytokine to a specific receptor molecule. Receptor binding then stimulates certain, and often independent, signal transduction pathways. (Kishimoto, T., et al., Cell 76:253-262 (1994).) The interaction between a cytokine and a receptor is a primary regulator of a variety of cellular processes, including activation, proliferation, and differentiation. (Arai, K. -I, et al., Ann. Rev. Biochem. 59:783-836 (1990); Paul, W. E. and Seder, R. A., Cell 76:241-251 (1994)).

Cytokines that bind to the interleukin-2 (IL-2) receptor common gamma chain (gamma c), including IL-2, IL-4, IL-7, IL-9, and IL-15, are important for the growth and differentiation of immune cells, such as T and B lymphocytes, natural killer cells, macrophages, and monocytes. These cytokines have overlapping biological effects that in part result from the use of the shared receptor subunit gamma c. Recently it has been shown that these cytokines activate a number of important intracellular signaling molecules, by binding to the interleukin-2 (IL-2) receptor common gamma chain (gamma c), including the Janus kinases JAK1 and JAK3 and members of the transcription factor family of signal transducers and activators of transcription (STATs).

The discovery of these signaling pathways has led to important new insights into their role in lymphocyte maturation, as it has emerged that mutations in the genes encoding both gamma c and JAK3 result in similar forms of severe combined immunodeficiency (SCID). For example, mutations in the human interleukin-2 (IL-2)

receptor gamma, mapped to the X chromosome, is associated with X-linked severe combined immunodeficiency (Human Molecular Genetics, 2(8): 1099 (1993).)

Thus, there is a need for polypeptides that regulate the differentiation and proliferation of cells, since disturbances of such regulation may be involved in disorders relating to immune system. Therefore, there is a need for identification and characterization of such human polypeptides which can play a role in detecting, preventing, ameliorating or correcting such disorders.

### *Summary of the Invention*

The present invention relates to a novel polynucleotide and the encoded polypeptide of CRCGCL. Moreover, the present invention relates to vectors, host cells, antibodies, and recombinant methods for producing the polypeptides and polynucleotides. Also provided are diagnostic methods for detecting disorders relates to the polypeptides, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying binding partners of CRCGCL.

### *Brief Description of the Drawings*

Figures 1A-1B show the nucleotide sequence (SEQ ID NO:1) and the deduced amino acid sequence (SEQ ID NO:2) of CRCGCL. The predicted leader sequence is located at about amino acids 1-22.

Figure 2 shows the regions of identity between the amino acid sequence of the CRCGCL protein and the translation product of the closest homolog, the Bos Taurus Interleukin-2 receptor gamma (Accession Nos. 1532088) (SEQ ID NO:3), determined by BLAST analysis. Identical amino acids between the two polypeptides are shaded in black, while conservative amino acids are boxed. By examining the regions of amino acids shaded and/or boxed, the skilled artisan can readily identify conserved domains between the two polypeptides. These conserved domains are preferred embodiments of the present invention.

Figure 3 shows an analysis of the CRCGCL amino acid sequence. Alpha, beta, turn and coil regions; hydrophilicity and hydrophobicity; amphipathic regions; flexible regions; antigenic index and surface probability are shown, and all were generated using the default settings. In the "Antigenic Index or Jameson-Wolf" graph, the positive peaks indicate locations of the highly antigenic regions of the CRCGCL



protein, i.e., regions from which epitope-bearing peptides of the invention can be obtained. The domains defined by these graphs are contemplated by the present invention. Tabular representation of the data summarized graphically in Figure 3 can be found in Table 1.

5

### *Detailed Description*

#### Definitions

The following definitions are provided to facilitate understanding of certain terms used throughout this specification.

10 In the present invention, "isolated" refers to material removed from its original environment (e.g., the natural environment if it is naturally occurring), and thus is altered "by the hand of man" from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be contained within a cell, and still be "isolated" because that vector, composition of  
15 matter, or particular cell is not the original environment of the polynucleotide.

In the present invention, a "secreted" CRCGCL protein refers to a protein capable of being directed to the ER, secretory vesicles, or the extracellular space as a result of a signal sequence, as well as a CRCGCL protein released into the extracellular space without necessarily containing a signal sequence. If the CRCGCL  
20 secreted protein is released into the extracellular space, the CRCGCL secreted protein can undergo extracellular processing to produce a "mature" CRCGCL protein. Release into the extracellular space can occur by many mechanisms, including exocytosis and proteolytic cleavage.

As used herein, a CRCGCL "polynucleotide" refers to a molecule having a  
25 nucleic acid sequence contained in SEQ ID NO:1 or the cDNA contained within the clone deposited with the ATCC. For example, the CRCGCL polynucleotide can contain the nucleotide sequence of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, with or without the signal sequence, the secreted protein coding region, as well as fragments, epitopes, domains, and  
30 variants of the nucleic acid sequence. Moreover, as used herein, a CRCGCL "polypeptide" refers to a molecule having the translated amino acid sequence generated from the polynucleotide as broadly defined. However, one embodiment of the present invention does not include the polynucleotide sequence of Genbank Accession No. X91553, herein incorporated by reference.

In specific embodiments, the polynucleotides of the invention are less than 300 kb, 200 kb, 100 kb, 50 kb, 15 kb, 10 kb, or 7.5 kb in length. In a further embodiment, polynucleotides of the invention comprise at least 15 contiguous nucleotides of CRCGCL coding sequence, but do not comprise all or a portion of any  
5 CRCGCL intron. In another embodiment, the nucleic acid comprising CRCGCL coding sequence does not contain coding sequences of a genomic flanking gene (i.e., 5' or 3' to the CRCGCL gene in the genome).

In the present invention, the full length CRCGCL sequence identified as SEQ ID NO:1 was generated by overlapping sequences of the deposited clone (contig  
10 analysis). A representative clone containing all or most of the sequence for SEQ ID NO:1 was deposited with the American Type Culture Collection ("ATCC") on March 23, 1998, and was given the ATCC Deposit Number 209691. A second clone was also deposited with the ATCC on February 25, 1998, and given ATCC Deposit Number 209641. The ATCC is located at 10801 University Boulevard, Manassas,  
15 VA 20110-2209, USA. The ATCC deposit was made pursuant to the terms of the Budapest Treaty on the international recognition of the deposit of microorganisms for purposes of patent procedure.

A CRCGCL "polynucleotide" also includes those polynucleotides capable of hybridizing, under stringent hybridization conditions, to sequences contained in SEQ  
20 ID NO:1, the complement thereof, or the cDNA within the deposited clone. "Stringent hybridization conditions" refers to an overnight incubation at 42 degree C in a solution comprising 50% formamide, 5x SSC (750 mM NaCl, 75 mM sodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 µg/ml denatured, sheared salmon sperm DNA, followed by washing  
25 the filters in 0.1x SSC at about 65 degree C.

Also contemplated are nucleic acid molecules that hybridize to the CRCGCL polynucleotides at moderately high stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of  
30 formamide result in lowered stringency); salt conditions, or temperature. For example, moderately high stringency conditions include an overnight incubation at 37 degree C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 0.2M NaH<sub>2</sub>PO<sub>4</sub>; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 µg/ml salmon sperm blocking DNA; followed by washes at 50 degree C with 1XSSPE, 0.1% SDS. In  
35 addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5X SSC).

Note that variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include Denhardt's reagent, BLOTTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

Of course, a polynucleotide which hybridizes only to polyA+ sequences (such as any 3' terminal polyA+ tract of a cDNA shown in the sequence listing), or to a complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

The CRCGCL polynucleotide can be composed of any polyribonucleotide or polydeoxiribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, CRCGCL polynucleotides can be composed of single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the CRCGCL polynucleotides can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. CRCGCL polynucleotides may also contain one or more modified bases or DNA or RNA backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications can be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

CRCGCL polypeptides can be composed of amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The CRCGCL polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in the CRCGCL polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given

CRCGCL polypeptide. Also, a given CRCGCL polypeptide may contain many types of modifications. CRCGCL polypeptides may be branched, for example, as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic CRCGCL polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. (See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifter et al., Meth Enzymol 182:626-646 (1990); Rattan et al., Ann NY Acad Sci 663:48-62 (1992).)

"SEQ ID NO:1" refers to a CRCGCL polynucleotide sequence while "SEQ ID NO:2" refers to a CRCGCL polypeptide sequence.

A CRCGCL polypeptide "having biological activity" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a CRCGCL polypeptide, including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be identical to that of the CRCGCL polypeptide, but rather substantially similar to the dose-dependence in a given activity as compared to the CRCGCL polypeptide (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity, and most preferably, not more than about three-fold less activity relative to the CRCGCL polypeptide.)

#### **CRCGCL Polynucleotides and Polypeptides**

Clone HTAEK53 was isolated from an activated T-cell cDNA library. Initially, the sequence of clone HTAEK53 was identified as SEQ ID NO:26 and the deduced amino acid sequence was predicted as SEQ ID NO:27, with a recognition that

### **CRCGCL Polynucleotides and Polypeptides**

Clone HTAEK53 was isolated from an activated T-cell cDNA library. Initially, the sequence of clone HTAEK53 was identified as SEQ ID NO:26 and the deduced amino acid sequence was predicted as SEQ ID NO:27, with a recognition that an apparent frame shift in the sequence existed between nucleotides 256 and 277. This frame shift was easily resolved using standard molecular biology techniques, generating the nucleotide sequence of SEQ ID NO:1 and the deduced amino acid sequence shown in SEQ ID NO:2.

The deposited clone contains a cDNA having a total of 1573 nucleotides, which encodes a predicted open reading frame of 371 amino acid residues. (See Figures 1A-1B.) The open reading frame begins at a N-terminal methionine located at nucleotide position 13, and ends at a stop codon at nucleotide position 1128. The predicted molecular weight of the CRCGCL protein should be about 42 kDa.

Subsequent Northern analysis also showed a 1.6 Kb transcript in a cervical cancer cell line (HeLa), activated T cells, and a lung carcinoma cell line (A549), while a shorter variant is also expressed in the lymph node and to a lesser extent in the spleen tissues, a pattern consistent with immune specific expression.

CRCGCL expression was not observed in the following cell lines, HL60, K562, Molt-4, Raji, SW480, G361, as well as the heart, brain, placenta, lung, liver, skeletal muscle, kidney, pancreas, thymus, prostate, testis, ovary, small intestine, colon, or peripheral blood leukocytes, a pattern consistent with immune specific expression.

Using BLAST analysis, SEQ ID NO:2 was found to be homologous to members of the Cytokine Receptor family. Particularly, SEQ ID NO:2 contains domains homologous to the translation product of the Bos Taurus mRNA for Interleukin-2 receptor gamma (Accession Nos. 1532088) (Figure 2) (SEQ ID NO:3), including the following conserved domains: (a) a predicted transmembrane domain located at about amino acids 226-260; (b) a predicted WXWS (SEQ ID NO:30), or [STGL]-x-W-[SG]-x-W-S (SEQ ID NO:18), domain located at about amino acids 198-204 (T-x-P-S-x-W-S) (SEQ ID NO:19), although not a perfect match; and (c) a predicted Jak Box, having the motif W(P,E)X(V,I)P(N,S,D)P (SEQ ID NO:20), domain located at about amino acids 261-268 (I-P-X-V-P-D-P) (SEQ ID NO:21), although not a perfect match. These polypeptide fragments of CRCGCL are specifically contemplated in the present invention. Because Interleukin-2 receptor gamma (Accession Nos. 1532088) is thought to be important as a cytokine receptor, the homology between Interleukin-2 receptor gamma (Accession Nos. 1532088) and

Moreover, the encoded polypeptide has a predicted leader sequence located at about amino acids 1-22. (See Figures 1A-1B.) Also shown in Figures 1A-1B, one embodiment of the secreted form of CRCGCL encompasses about amino acids 23-371, amino acids 23-225, or amino acids 1-231. These polypeptide fragments of  
5 CRCGCL are specifically contemplated in the present invention.

Other preferred polypeptide fragments comprise the amino acids sequence: QIQHYFNLETVQVTWNASKYSRTNLTFFHYRFNGDEAYDQCTNYLLQEGHITSG C (SEQ ID NO:22); RRHSLFLHQEWDAPRFHIRKSLDGLLPETQF (SEQ ID NO:23);  
10 LLYEVQYRSPFDTEWQSKQENTCNVTIEGLDAEKCYSEFWVRVKAMEDVYGPDTYPSDWSEVTCWQRGEIRDACAETPTPPK (SEQ ID NO:24); and/or MEDVYGPDTYPSDWSEVTCWQRGEIRDACAETPTPPKPKLSKFILISSLAILLM VLLLLLSLWKLWRXKKFLXPSVPDPKSIFPGLFXIHQGNFQEWITDTQNVAIH LHKMAGAEQESGPEEPLVVQLAKTEAESPRMLDPQTEEKEASGGSLQLPHQP  
15 LQGGDVVTIGGFTFVMNDRSYVA (SEQ ID NO:25), as well as fragments thereof. Also preferred are polynucleotide fragments encoding these polypeptide fragments.

Because CRCGCL was isolated from activated T cells, nucleic acids of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of immune disorders.  
20 Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the immune system, expression of this gene at significantly higher or lower levels may be detected in certain tissues (e.g., cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or  
25 spinal fluid) taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue from an individual not having the disorder.

The tissue distribution in only activated T-cells and homology to the cytokine receptors IL2 and IL13 suggests that this protein is a novel member of the cytokine  
30 receptor family expressed specifically on T-cells. The tissue distribution of this gene in cells of the immune system suggests that the protein product of this clone would be useful for treatment, prophylaxis and diagnosis of immune and autoimmune diseases, such as lupus, transplant rejection, allergic reactions, arthritis, asthma, immunodeficiency diseases, leukemia, AIDS. In addition its expression in T-cells  
35 suggests a potential role in the treatment, prophylaxis and detection of thymus disorders such as Graves Disease, lymphocytic thyroiditis, hyperthyroidism and hypothyroidism. The receptor could also serve as a target for small molecule or

monoclonal antibody, blocking its activity, which could be important in the disease states listed herein.

5 The CRGCL nucleotide sequence identified as SEQ ID NO:1 was assembled from partially homologous ("overlapping") sequences obtained from the deposited clone, and in some cases, from additional related DNA clones. The overlapping sequences were assembled into a single contiguous sequence of high redundancy (usually three to five overlapping sequences at each nucleotide position), resulting in a final sequence identified as SEQ ID NO:1.

10 Therefore, SEQ ID NO:1 and the translated SEQ ID NO:2 are sufficiently accurate and otherwise suitable for a variety of uses well known in the art and described further below. For instance, SEQ ID NO:1 is useful for designing nucleic acid hybridization probes that will detect nucleic acid sequences contained in SEQ ID NO:1 or the cDNA contained in the deposited clone. These probes will also hybridize to nucleic acid molecules in biological samples, thereby enabling a variety of forensic and diagnostic methods of the invention. Similarly, polypeptides identified from SEQ ID NO:2 may be used to generate antibodies which bind specifically to CRGCL.

15 Nevertheless, DNA sequences generated by sequencing reactions can contain sequencing errors. The errors exist as misidentified nucleotides, or as insertions or deletions of nucleotides in the generated DNA sequence. The erroneously inserted or deleted nucleotides cause frame shifts in the reading frames of the predicted amino acid sequence. In these cases, the predicted amino acid sequence diverges from the actual amino acid sequence, even though the generated DNA sequence may be greater than 99.9% identical to the actual DNA sequence (for example, one base insertion or deletion in an open reading frame of over 1000 bases).

25 Accordingly, for those applications requiring precision in the nucleotide sequence or the amino acid sequence, the present invention provides not only the generated nucleotide sequence identified as SEQ ID NO:1 and the predicted translated amino acid sequence identified as SEQ ID NO:2, but also a sample of plasmid DNA containing a human cDNA of CRGCL deposited with the ATCC. The nucleotide sequence of the deposited CRGCL clone can readily be determined by sequencing the deposited clone in accordance with known methods. The predicted CRGCL amino acid sequence can then be verified from such deposits. Moreover, the amino acid sequence of the protein encoded by the deposited clone can also be directly determined by peptide sequencing or by expressing the protein in a suitable host cell containing the deposited human CRGCL cDNA, collecting the protein, and  
35 determining its sequence.

The present invention also relates to the CRCGCL gene corresponding to SEQ ID NO:1, SEQ ID NO:2, or the deposited clone. The CRCGCL gene can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include preparing probes or primers from the disclosed sequence and  
5 identifying or amplifying the CRCGCL gene from appropriate sources of genomic material.

Also provided in the present invention are species homologs of CRCGCL. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source for  
10 the desired homologue.

The CRCGCL polypeptides can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides  
15 are well understood in the art.

The CRCGCL polypeptides may be in the form of the secreted protein, including the mature form, or may be a part of a larger protein, such as a fusion protein (see below). It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences  
20 which aid in purification, such as multiple histidine residues, or an additional sequence for stability during recombinant production.

CRCGCL polypeptides are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of a CRCGCL polypeptide, including the secreted polypeptide, can be substantially  
25 purified by the one-step method described in Smith and Johnson, Gene 67:31-40 (1988). CRCGCL polypeptides also can be purified from natural or recombinant sources using antibodies of the invention raised against the CRCGCL protein in methods which are well known in the art.

### 30 **Polynucleotide and Polypeptide Variants**

"Variant" refers to a polynucleotide or polypeptide differing from the CRCGCL polynucleotide or polypeptide, but retaining essential properties thereof. Generally, variants are overall closely similar, and, in many regions, identical to the CRCGCL polynucleotide or polypeptide.

35 By a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence of the present invention, it is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence



except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the CRGGL polypeptide. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. The query sequence may be an entire sequence shown of SEQ ID NO:1, the ORF (open reading frame), or any fragment specified as described herein.

As a practical matter, whether any particular nucleic acid molecule or polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to a nucleotide sequence of the present invention can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245.) In a sequence alignment the query and subject sequences are both DNA sequences. An RNA sequence can be compared by converting U's to T's. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB alignment of DNA sequences to calculate percent identity are: Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty 0.05, Window Size=500 or the length of the subject nucleotide sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence because of 5' or 3' deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for 5' and 3' truncations of the subject sequence when calculating percent identity. For subject sequences truncated at the 5' or 3' ends, relative to the query sequence, the percent identity is corrected by calculating the number of bases of the query sequence that are 5' and 3' of the subject sequence, which are not matched/aligned, as a percent of the total bases of the query sequence. Whether a nucleotide is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This corrected score is what is used for the purposes of the present invention. Only bases outside the 5' and 3' bases of the subject sequence, as displayed by the FASTDB alignment,

which are not matched/aligned with the query sequence, are calculated for the purposes of manually adjusting the percent identity score.

For example, a 90 base subject sequence is aligned to a 100 base query sequence to determine percent identity. The deletions occur at the 5' end of the subject sequence and therefore, the FASTDB alignment does not show a matched/alignment of the first 10 bases at 5' end. The 10 unpaired bases represent 10% of the sequence (number of bases at the 5' and 3' ends not matched/total number of bases in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 bases were perfectly matched the final percent identity would be 90%. In another example, a 90 base subject sequence is compared with a 100 base query sequence. This time the deletions are internal deletions so that there are no bases on the 5' or 3' of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only bases 5' and 3' of the subject sequence which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to made for the purposes of the present invention.

By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a query amino acid sequence of the present invention, it is intended that the amino acid sequence of the subject polypeptide is identical to the query sequence except that the subject polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the query amino acid sequence. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a query amino acid sequence, up to 5% of the amino acid residues in the subject sequence may be inserted, deleted, (indels) or substituted with another amino acid. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the amino acid sequences shown in SEQ ID NO:2 or to the amino acid sequence encoded by deposited DNA clone can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence

alignment the query and subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization  
5 Group Length=0, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject amino acid sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence due to N- or C-terminal deletions, not because of internal deletions, a manual correction must be made  
10 to the results. This is because the FASTDB program does not account for N- and C-terminal truncations of the subject sequence when calculating global percent identity. For subject sequences truncated at the N- and C-termini, relative to the query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not  
15 matched/aligned with a corresponding subject residue, as a percent of the total bases of the query sequence. Whether a residue is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score  
20 is what is used for the purposes of the present invention. Only residues to the N- and C-termini of the subject sequence, which are not matched/aligned with the query sequence, are considered for the purposes of manually adjusting the percent identity score. That is, only query residue positions outside the farthest N- and C-terminal residues of the subject sequence.

25 For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the N-terminus of the subject sequence and therefore, the FASTDB alignment does not show a matching/alignment of the first 10 residues at the N-terminus. The 10 unpaired residues represent 10% of the sequence (number of residues at the N- and C-termini  
30 not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another example, a 90 residue subject sequence is compared with a 100 residue query sequence. This time the deletions are internal deletions so there are no residues at the  
35 N- or C-termini of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject

sequence, as displayed in the FASTDB alignment, which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to made for the purposes of the present invention.

The CRCGCL variants may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing alterations which produce silent substitutions, additions, or deletions, but do not alter the properties or activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred. Moreover, variants in which 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination are also preferred. CRCGCL polynucleotide variants can be produced for a variety of reasons, e.g., to optimize codon expression for a particular host (change codons in the human mRNA to those preferred by a bacterial host such as *E. coli*).

Naturally occurring CRCGCL variants are called "allelic variants," and refer to one of several alternate forms of a gene occupying a given locus on a chromosome of an organism. (Genes II, Lewin, B., ed., John Wiley & Sons, New York (1985).) These allelic variants can vary at either the polynucleotide and/or polypeptide level. Alternatively, non-naturally occurring variants may be produced by mutagenesis techniques or by direct synthesis.

Using known methods of protein engineering and recombinant DNA technology, variants may be generated to improve or alter the characteristics of the CRCGCL polypeptides. For instance, one or more amino acids can be deleted from the N-terminus or C-terminus of the secreted protein without substantial loss of biological function. The authors of Ron et al., *J. Biol. Chem.* 268: 2984-2988 (1993), reported variant KGF proteins having heparin binding activity even after deleting 3, 8, or 27 amino-terminal amino acid residues. Similarly, Interferon gamma exhibited up to ten times higher activity after deleting 8-10 amino acid residues from the carboxy terminus of this protein. (Dobeli et al., *J. Biotechnology* 7:199-216 (1988).)

Moreover, ample evidence demonstrates that variants often retain a biological activity similar to that of the naturally occurring protein. For example, Gayle and coworkers (*J. Biol. Chem.* 268:22105-22111 (1993)) conducted extensive mutational analysis of human cytokine IL-1a. They used random mutagenesis to generate over 3,500 individual IL-1a mutants that averaged 2.5 amino acid changes per variant over the entire length of the molecule. Multiple mutations were examined at every possible amino acid position. The investigators found that "[m]ost of the molecule could be altered with little effect on either [binding or biological activity]." (See, Abstract.) In

fact, only 23 unique amino acid sequences, out of more than 3,500 nucleotide sequences examined, produced a protein that significantly differed in activity from wild-type.

Furthermore, even if deleting one or more amino acids from the N-terminus or C-terminus of a polypeptide results in modification or loss of one or more biological functions, other biological activities may still be retained. For example, the ability of a deletion variant to induce and/or to bind antibodies which recognize the secreted form will likely be retained when less than the majority of the residues of the secreted form are removed from the N-terminus or C-terminus. Whether a particular polypeptide lacking N- or C-terminal residues of a protein retains such immunogenic activities can readily be determined by routine methods described herein and otherwise known in the art.

Thus, the invention further includes CRGCL polypeptide variants which show substantial biological activity. Such variants include deletions, insertions, inversions, repeats, and substitutions selected according to general rules known in the art so as to have little effect on activity. For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie, J. U. et al., *Science* 247:1306-1310 (1990), wherein the authors indicate that there are two main strategies for studying the tolerance of an amino acid sequence to change.

The first strategy exploits the tolerance of amino acid substitutions by natural selection during the process of evolution. By comparing amino acid sequences in different species, conserved amino acids can be identified. These conserved amino acids are likely important for protein function. In contrast, the amino acid positions where substitutions have been tolerated by natural selection indicates that these positions are not critical for protein function. Thus, positions tolerating amino acid substitution could be modified while still maintaining biological activity of the protein.

The second strategy uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene to identify regions critical for protein function. For example, site directed mutagenesis or alanine-scanning mutagenesis (introduction of single alanine mutations at every residue in the molecule) can be used. (Cunningham and Wells, *Science* 244:1081-1085 (1989).) The resulting mutant molecules can then be tested for biological activity.

As the authors state, these two strategies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at certain amino acid positions in the protein. For example, most buried (within the tertiary structure of the protein) amino acid residues require nonpolar side chains, whereas few features of surface side

chains are generally conserved. Moreover, tolerated conservative amino acid substitutions involve replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues Asp and Glu; replacement of the amide residues Asn and Gln; replacement of the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr, and Trp, and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly.

Besides conservative amino acid substitution, variants of CRCGCL include (i) substitutions with one or more of the non-conserved amino acid residues, where the substituted amino acid residues may or may not be one encoded by the genetic code, or (ii) substitution with one or more of amino acid residues having a substituent group, or (iii) fusion of the mature polypeptide with another compound, such as a compound to increase the stability and/or solubility of the polypeptide (for example, polyethylene glycol), or (iv) fusion of the polypeptide with additional amino acids, such as an IgG Fc fusion region peptide, or leader or secretory sequence, or a sequence facilitating purification. Such variant polypeptides are deemed to be within the scope of those skilled in the art from the teachings herein.

For example, CRCGCL polypeptide variants containing amino acid substitutions of charged amino acids with other charged or neutral amino acids may produce proteins with improved characteristics, such as less aggregation. Aggregation of pharmaceutical formulations both reduces activity and increases clearance due to the aggregate's immunogenic activity. (Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967); Robbins et al., Diabetes 36: 838-845 (1987); Cleland et al., Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377 (1993).)

A further embodiment of the invention relates to a polypeptide which comprises the amino acid sequence of a CRCGCL polypeptide having an amino acid sequence which contains at least one amino acid substitution, but not more than 50 amino acid substitutions, even more preferably, not more than 40 amino acid substitutions, still more preferably, not more than 30 amino acid substitutions, and still even more preferably, not more than 20 amino acid substitutions. Of course, in order of ever-increasing preference, it is highly preferable for a peptide or polypeptide to have an amino acid sequence which comprises the amino acid sequence of a CRCGCL polypeptide, which contains at least one, but not more than 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1 amino acid substitutions. In specific embodiments, the number of additions, substitutions, and/or deletions in the amino acid sequence of Figures 1A-1B or fragments thereof (e.g., the mature form and/or other fragments described herein),

is 1-5, 5-10, 5-25, 5-50, 10-50 or 50-150, conservative amino acid substitutions are preferable.

### **Polynucleotide and Polypeptide Fragments**

5 In the present invention, a "polynucleotide fragment" refers to a short polynucleotide having a nucleic acid sequence contained in the deposited clone or shown in SEQ ID NO:1. The short nucleotide fragments are preferably at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably, at least about 40 nt in length. A fragment "at least 20 nt  
10 in length," for example, is intended to include 20 or more contiguous bases from the cDNA sequence contained in the deposited clone or the nucleotide sequence shown in SEQ ID NO:1. These nucleotide fragments are useful as diagnostic probes and primers as discussed herein. Of course, larger fragments (e.g., 50, 150, 500, 600, 2000 nucleotides) are preferred.

15 Moreover, representative examples of CRCGCL polynucleotide fragments include, for example, fragments having a sequence from about nucleotide number 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, 451-500, 501-550, 551-600, 651-700, 701-750, 751-800, 800-850, 851-900, 901-950, 951-1000, 1001-1050, 1051-1100, 1101-1150, 1151-1200, 1201-1250, 1251-  
20 1300, 1301-1350, 1351-1400, 1401-1450, 1451-1500, 1501-1550, 1551-1600, 1601-1650, 1651-1700, 1701-1750, 1751-1800, 1801-1850, 1851-1900, 1901-1950, 1951-2000, or 2001 to the end of SEQ ID NO:1 or the cDNA contained in the deposited clone. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both  
25 termini. Preferably, these fragments encode a polypeptide which has biological activity. More preferably, these polynucleotides can be used as probes or primers as discussed herein.

In the present invention, a "polypeptide fragment" refers to a short amino acid sequence contained in SEQ ID NO:2 or encoded by the cDNA contained in the  
30 deposited clone. Protein fragments may be "free-standing," or comprised within a larger polypeptide of which the fragment forms a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, 102-120, 121-140, 141-160, 161-180, 181-200, 201-  
35 220, 221-240, 241-260, 261-280, or 281 to the end of the coding region. Moreover, polypeptide fragments can be about 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, or 150 amino acids in length. In this context "about" includes the

particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) amino acids, at either extreme or at both extremes.

Preferred polypeptide fragments include the secreted CRCGCL protein as well as the mature form. Further preferred polypeptide fragments include the secreted  
 5 CRCGCL protein or the mature form having a continuous series of deleted residues from the amino or the carboxy terminus, or both. For example, any number of amino acids, ranging from 1-60, can be deleted from the amino terminus of either the secreted CRCGCL polypeptide or the mature form. Similarly, any number of amino acids, ranging from 1-30, can be deleted from the carboxy terminus of the secreted  
 10 CRCGCL protein or mature form. Furthermore, any combination of the above amino and carboxy terminus deletions are preferred. Similarly, polynucleotide fragments encoding these CRCGCL polypeptide fragments are also preferred.

Particularly, N-terminal deletions of the CRCGCL polypeptide can be described by the general formula m-371, where m is an integer from 2 to 370, where  
 15 m corresponds to the position of the amino acid residue identified in SEQ ID NO:2. More in particular, the invention provides polynucleotides encoding polypeptides comprising, or alternatively consisting of, the amino acid sequence of residues of N-terminal deletions of the CRCGCL polypeptide of the invention shown as SEQ ID NO:2 include polypeptides comprising the amino acid sequence of residues: G-2 to L-  
 20 371; R-3 to L-371; L-4 to L-371; V-5 to L-371; L-6 to L-371; L-7 to L-371; W-8 to L-371; G-9 to L-371; A-10 to L-371; A-11 to L-371; V-12 to L-371; F-13 to L-371; L-14 to L-371; L-15 to L-371; G-16 to L-371; G-17 to L-371; W-18 to L-371; M-19 to L-371; A-20 to L-371; L-21 to L-371; G-22 to L-371; Q-23 to L-371; G-24 to L-371; G-25 to L-371; A-26 to L-371; A-27 to L-371; E-28 to L-371; G-29 to L-371; V-30 to  
 25 L-371; Q-31 to L-371; I-32 to L-371; Q-33 to L-371; I-34 to L-371; I-35 to L-371; Y-36 to L-371; F-37 to L-371; N-38 to L-371; L-39 to L-371; E-40 to L-371; T-41 to L-371; V-42 to L-371; Q-43 to L-371; V-44 to L-371; T-45 to L-371; W-46 to L-371; N-47 to L-371; A-48 to L-371; S-49 to L-371; K-50 to L-371; Y-51 to L-371; S-52 to L-371; R-53 to L-371; T-54 to L-371; N-55 to L-371; L-56 to L-371; T-57 to L-371;  
 30 F-58 to L-371; H-59 to L-371; Y-60 to L-371; R-61 to L-371; F-62 to L-371; N-63 to L-371; G-64 to L-371; D-65 to L-371; E-66 to L-371; A-67 to L-371; Y-68 to L-371; D-69 to L-371; Q-70 to L-371; C-71 to L-371; T-72 to L-371; N-73 to L-371; Y-74 to L-371; L-75 to L-371; L-76 to L-371; Q-77 to L-371; E-78 to L-371; G-79 to L-371; H-80 to L-371; T-81 to L-371; S-82 to L-371; G-83 to L-371; C-84 to L-371; L-85 to  
 35 L-371; L-86 to L-371; D-87 to L-371; A-88 to L-371; E-89 to L-371; Q-90 to L-371; R-91 to L-371; D-92 to L-371; D-93 to L-371; I-94 to L-371; L-95 to L-371; Y-96 to L-371; F-97 to L-371; S-98 to L-371; I-99 to L-371; R-100 to L-371; N-101 to L-



R-91 to L-371; D-92 to L-371; D-93 to L-371; I-94 to L-371; L-95 to L-371; Y-96 to L-371; F-97 to L-371; S-98 to L-371; I-99 to L-371; R-100 to L-371; N-101 to L-371; G-102 to L-371; T-103 to L-371; H-104 to L-371; P-105 to L-371; V-106 to L-371; F-107 to L-371; T-108 to L-371; A-109 to L-371; S-110 to L-371; R-111 to L-371; W-112 to L-371; M-113 to L-371; V-114 to L-371; Y-115 to L-371; Y-116 to L-371; L-117 to L-371; K-118 to L-371; P-119 to L-371; S-120 to L-371; S-121 to L-371; P-122 to L-371; K-123 to L-371; H-124 to L-371; V-125 to L-371; R-126 to L-371; F-127 to L-371; S-128 to L-371; W-129 to L-371; H-130 to L-371; Q-131 to L-371; D-132 to L-371; A-133 to L-371; V-134 to L-371; T-135 to L-371; V-136 to L-371; T-137 to L-371; C-138 to L-371; S-139 to L-371; D-140 to L-371; L-141 to L-371; S-142 to L-371; Y-143 to L-371; G-144 to L-371; D-145 to L-371; L-146 to L-371; L-147 to L-371; Y-148 to L-371; E-149 to L-371; V-150 to L-371; Q-151 to L-371; Y-152 to L-371; R-153 to L-371; S-154 to L-371; P-155 to L-371; F-156 to L-371; D-157 to L-371; T-158 to L-371; E-159 to L-371; W-160 to L-371; Q-161 to L-371; S-162 to L-371; K-163 to L-371; Q-164 to L-371; E-165 to L-371; N-166 to L-371; T-167 to L-371; C-168 to L-371; N-169 to L-371; V-170 to L-371; F-171 to L-371; I-172 to L-371; E-173 to L-371; G-174 to L-371; L-175 to L-371; D-176 to L-371; A-177 to L-371; E-178 to L-371; K-179 to L-371; C-180 to L-371; Y-181 to L-371; S-182 to L-371; F-183 to L-371; W-184 to L-371; V-185 to L-371; R-186 to L-371; V-187 to L-371; K-188 to L-371; A-189 to L-371; M-190 to L-371; E-191 to L-371; D-192 to L-371; V-193 to L-371; Y-194 to L-371; G-195 to L-371; P-196 to L-371; D-197 to L-371; T-198 to L-371; Y-199 to L-371; P-200 to L-371; S-201 to L-371; D-202 to L-371; W-203 to L-371; S-204 to L-371; E-205 to L-371; V-206 to L-371; T-207 to L-371; C-208 to L-371; W-209 to L-371; Q-210 to L-371; R-211 to L-371; G-212 to L-371; E-213 to L-371; I-214 to L-371; R-215 to L-371; D-216 to L-371; A-217 to L-371; C-218 to L-371; A-219 to L-371; E-220 to L-371; T-221 to L-371; P-222 to L-371; T-223 to L-371; P-224 to L-371; P-225 to L-371; K-226 to L-371; P-227 to L-371; K-228 to L-371; L-229 to L-371; S-230 to L-371; K-231 to L-371; F-232 to L-371; I-233 to L-371; L-234 to L-371; I-235 to L-371; S-236 to L-371; S-237 to L-371; L-238 to L-371; A-239 to L-371; I-240 to L-371; L-241 to L-371; L-242 to L-371; M-243 to L-371; V-244 to L-371; S-245 to L-371; L-246 to L-371; L-247 to L-371; L-248 to L-371; L-249 to L-371; S-250 to L-371; L-251 to L-371; W-252 to L-371; K-253 to L-371; L-254 to L-371; W-255 to L-371; R-256 to L-371; V-257 to L-371; K-258 to L-371; K-259 to L-371; F-260 to L-371; L-261 to L-371; I-262 to L-371; P-263 to L-371; S-264 to L-371; V-265 to L-371; P-266 to L-371; D-267 to L-371; P-268 to L-371; K-269 to L-371; S-270 to L-371; I-271 to L-371; F-272 to L-371; P-273 to L-371; G-274 to L-371; L-275 to L-371; F-276 to L-

371; E-277 to L-371; I-278 to L-371; H-279 to L-371; Q-280 to L-371; G-281 to L-371; N-282 to L-371; F-283 to L-371; Q-284 to L-371; E-285 to L-371; W-286 to L-371; I-287 to L-371; F-288 to L-371; D-289 to L-371; T-290 to L-371; Q-291 to L-371; N-292 to L-371; V-293 to L-371; A-294 to L-371; H-295 to L-371; L-296 to L-371; H-297 to L-371; K-298 to L-371; M-299 to L-371; A-300 to L-371; G-301 to L-371; A-302 to L-371; E-303 to L-371; Q-304 to L-371; E-305 to L-371; S-306 to L-371; G-307 to L-371; P-308 to L-371; E-309 to L-371; E-310 to L-371; P-311 to L-371; L-312 to L-371; V-313 to L-371; V-314 to L-371; Q-315 to L-371; L-316 to L-371; A-317 to L-371; K-318 to L-371; T-319 to L-371; E-320 to L-371; A-321 to L-371; E-322 to L-371; S-323 to L-371; P-324 to L-371; R-325 to L-371; M-326 to L-371; L-327 to L-371; D-328 to L-371; P-329 to L-371; Q-330 to L-371; T-331 to L-371; E-332 to L-371; E-333 to L-371; K-334 to L-371; E-335 to L-371; A-336 to L-371; S-337 to L-371; G-338 to L-371; G-339 to L-371; S-340 to L-371; L-341 to L-371; Q-342 to L-371; L-343 to L-371; P-344 to L-371; H-345 to L-371; Q-346 to L-371; P-347 to L-371; L-348 to L-371; Q-349 to L-371; G-350 to L-371; G-351 to L-371; D-352 to L-371; V-353 to L-371; V-354 to L-371; T-355 to L-371; I-356 to L-371; G-357 to L-371; G-358 to L-371; F-359 to L-371; T-360 to L-371; F-361 to L-371; V-362 to L-371; M-363 to L-371; N-364 to L-371; D-365 to L-371; R-366 to L-371; of SEQ ID NO:2. Polynucleotides encoding these polypeptides are also encompassed by the invention.

Moreover, C-terminal deletions of the CRCGCL polypeptide can also be described by the general formula 1 n, where n is an integer from 2 to 371, where n corresponds to the position of amino acid residue identified in SEQ ID NO:2. More in particular, the invention provides polynucleotides encoding polypeptides comprising, or alternatively consisting of, the amino acid sequence of residues of C-terminal deletions of the CRCGCL polypeptide of the invention shown as SEQ ID NO:2 include polypeptides comprising the amino acid sequence of residues: M-1 to A-370; M-1 to V-369; M-1 to Y-368; M-1 to S-367; M-1 to R-366; M-1 to D-365; M-1 to N-364; M-1 to M-363; M-1 to V-362; M-1 to F-361; M-1 to T-360; M-1 to F-359; M-1 to G-358; M-1 to G-357; M-1 to I-356; M-1 to T-355; M-1 to V-354; M-1 to V-353; M-1 to D-352; M-1 to G-351; M-1 to G-350; M-1 to Q-349; M-1 to L-348; M-1 to P-347; M-1 to Q-346; M-1 to H-345; M-1 to P-344; M-1 to L-343; M-1 to Q-342; M-1 to L-341; M-1 to S-340; M-1 to G-339; M-1 to G-338; M-1 to S-337; M-1 to A-336; M-1 to E-335; M-1 to K-334; M-1 to E-333; M-1 to E-332; M-1 to T-331; M-1 to Q-330; M-1 to P-329; M-1 to D-328; M-1 to L-327; M-1 to M-326; M-1 to R-325; M-1 to P-324; M-1 to S-323; M-1 to E-322; M-1 to A-321; M-1 to E-320; M-1 to T-319; M-1 to K-318; M-1 to A-317; M-1 to L-316; M-1 to Q-315; M-1 to V-314; M-1 to V-

l to G-301; M-1 to A-300; M-1 to M-299; M-1 to K-298; M-1 to H-297; M-1 to L-296; M-1 to H-295; M-1 to A-294; M-1 to V-293; M-1 to N-292; M-1 to Q-291; M-1 to I-290; M-1 to D-289; M-1 to T-288; M-1 to I-287; M-1 to W-286; M-1 to E-285; M-1 to Q-284; M-1 to F-283; M-1 to N-282; M-1 to G-281; M-1 to Q-280; M-1 to H-279; M-1 to I-278; M-1 to E-277; M-1 to F-276; M-1 to L-275; M-1 to G-274; M-1 to P-273; M-1 to F-272; M-1 to I-271; M-1 to S-270; M-1 to K-269; M-1 to P-268; M-1 to D-267; M-1 to P-266; M-1 to V-265; M-1 to S-264; M-1 to P-263; M-1 to I-262; M-1 to L-261; M-1 to F-260; M-1 to K-259; M-1 to K-258; M-1 to V-257; M-1 to R-256; M-1 to W-255; M-1 to L-254; M-1 to K-253; M-1 to W-252; M-1 to L-251; M-1 to S-250; M-1 to L-249; M-1 to L-248; M-1 to L-247; M-1 to L-246; M-1 to S-245; M-1 to V-244; M-1 to M-243; M-1 to L-242; M-1 to L-241; M-1 to I-240; M-1 to A-239; M-1 to L-238; M-1 to S-237; M-1 to S-236; M-1 to I-235; M-1 to L-234; M-1 to I-233; M-1 to F-232; M-1 to K-231; M-1 to S-230; M-1 to L-229; M-1 to K-228; M-1 to P-227; M-1 to K-226; M-1 to P-225; M-1 to P-224; M-1 to T-223; M-1 to P-222; M-1 to T-221; M-1 to E-220; M-1 to A-219; M-1 to C-218; M-1 to A-217; M-1 to D-216; M-1 to R-215; M-1 to I-214; M-1 to E-213; M-1 to G-212; M-1 to R-211; M-1 to Q-210; M-1 to W-209; M-1 to C-208; M-1 to T-207; M-1 to V-206; M-1 to E-205; M-1 to S-204; M-1 to W-203; M-1 to D-202; M-1 to S-201; M-1 to P-200; M-1 to Y-199; M-1 to T-198; M-1 to D-197; M-1 to P-196; M-1 to G-195; M-1 to Y-194; M-1 to V-193; M-1 to D-192; M-1 to E-191; M-1 to M-190; M-1 to A-189; M-1 to K-188; M-1 to V-187; M-1 to R-186; M-1 to V-185; M-1 to W-184; M-1 to F-183; M-1 to S-182; M-1 to Y-181; M-1 to C-180; M-1 to K-179; M-1 to E-178; M-1 to A-177; M-1 to D-176; M-1 to L-175; M-1 to G-174; M-1 to E-173; M-1 to I-172; M-1 to T-171; M-1 to V-170; M-1 to N-169; M-1 to C-168; M-1 to T-167; M-1 to N-166; M-1 to E-165; M-1 to Q-164; M-1 to K-163; M-1 to S-162; M-1 to Q-161; M-1 to W-160; M-1 to E-159; M-1 to T-158; M-1 to D-157; M-1 to F-156; M-1 to P-155; M-1 to S-154; M-1 to R-153; M-1 to Y-152; M-1 to Q-151; M-1 to V-150; M-1 to E-149; M-1 to Y-148; M-1 to L-147; M-1 to L-146; M-1 to D-145; M-1 to G-144; M-1 to Y-143; M-1 to S-142; M-1 to L-141; M-1 to D-140; M-1 to S-139; M-1 to C-138; M-1 to T-137; M-1 to V-136; M-1 to T-135; M-1 to V-134; M-1 to A-133; M-1 to D-132; M-1 to Q-131; M-1 to H-130; M-1 to W-129; M-1 to S-128; M-1 to F-127; M-1 to R-126; M-1 to V-125; M-1 to H-124; M-1 to K-123; M-1 to P-122; M-1 to S-121; M-1 to S-120; M-1 to P-119; M-1 to K-118; M-1 to L-117; M-1 to Y-116; M-1 to Y-115; M-1 to V-114; M-1 to M-113; M-1 to W-112; M-1 to R-111; M-1 to S-110; M-1 to A-109; M-1 to T-108; M-1 to F-107; M-1 to V-106; M-1 to P-105; M-1 to H-104; M-1 to T-103; M-1 to G-102; M-1 to N-101; M-1 to R-100; M-1 to I-99; M-1 to S-98; M-1 to F-97; M-1 to Y-96; M-1 to L-95; M-1 to I-94; M-1 to D-93; M-1 to D-92; M-1 to R-91; M-1

to Q-90; M-1 to E-89; M-1 to A-88; M-1 to D-87; M-1 to L-86; M-1 to L-85; M-1 to C-84; M-1 to G-83; M-1 to S-82; M-1 to T-81; M-1 to H-80; M-1 to G-79; M-1 to E-78; M-1 to Q-77; M-1 to L-76; M-1 to L-75; M-1 to Y-74; M-1 to N-73; M-1 to T-72; M-1 to C-71; M-1 to Q-70; M-1 to D-69; M-1 to Y-68; M-1 to A-67; M-1 to E-66; M-1 to D-65; M-1 to G-64; M-1 to N-63; M-1 to F-62; M-1 to R-61; M-1 to Y-60; M-1 to H-59; M-1 to F-58; M-1 to T-57; M-1 to L-56; M-1 to N-55; M-1 to T-54; M-1 to R-53; M-1 to S-52; M-1 to Y-51; M-1 to K-50; M-1 to S-49; M-1 to A-48; M-1 to N-47; M-1 to W-46; M-1 to T-45; M-1 to V-44; M-1 to Q-43; M-1 to V-42; M-1 to T-41; M-1 to E-40; M-1 to L-39; M-1 to N-38; M-1 to F-37; M-1 to Y-36; M-1 to I-35; M-1 to I-34; M-1 to Q-33; M-1 to I-32; M-1 to Q-31; M-1 to V-30; M-1 to G-29; M-1 to E-28; M-1 to A-27; M-1 to A-26; M-1 to G-25; M-1 to G-24; M-1 to Q-23; M-1 to G-22; M-1 to L-21; M-1 to A-20; M-1 to M-19; M-1 to W-18; M-1 to G-17; M-1 to G-16; M-1 to L-15; M-1 to L-14; M-1 to F-13; M-1 to V-12; M-1 to A-11; M-1 to A-10; M-1 to G-9; M-1 to W-8; M-1 to L-7; of SEQ ID NO:2. Polynucleotides encoding these polypeptides are also encompassed by the invention.

In addition, any of the above listed N- or C-terminal deletions can be combined to produce a N- and C-terminal deleted CRCGCL polypeptide. The invention also provides polypeptides having one or more amino acids deleted from both the amino and the carboxyl termini, which may be described generally as having residues m-n of SEQ ID NO:2, where n and m are integers as described above. Polynucleotides encoding these polypeptides are also encompassed by the invention.

Also preferred are CRCGCL polypeptide and polynucleotide fragments characterized by structural or functional domains. Preferred embodiments of the invention include fragments that comprise alpha-helix and alpha-helix forming regions ("alpha-regions"), beta-sheet and beta-sheet-forming regions ("beta-regions"), turn and turn-forming regions ("turn-regions"), coil and coil-forming regions ("coil-regions"), hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions. As set out in the Figures, such preferred regions include Garnier-Robson alpha-regions, beta-regions, turn-regions, and coil-regions, Chou-Fasman alpha-regions, beta-regions, and turn-regions, Kyte-Doolittle hydrophilic regions and hydrophobic regions, Eisenberg alpha and beta amphipathic regions, Karplus-Schulz flexible regions, Emini surface-forming regions, and Jameson-Wolf high antigenic index regions. Polypeptide fragments of SEQ ID NO:2 falling within conserved domains are specifically contemplated by the present invention. (See Figure 3 and Table 1.) Moreover, polynucleotide fragments encoding these domains are also contemplated.

Other preferred fragments are biologically active CRCGCL fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the CRCGCL polypeptide. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity.

However, many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:1 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. One embodiment of the present invention excludes Genbank Accession No. X91553 (herein incorporated by reference in its entirety.) Moreover, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1559 of SEQ ID NO:1, b is an integer of 15 to 1573, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:1, and where the b is greater than or equal to a + 14.

#### **Epitope-Bearing Portions**

In another aspect, the invention provides peptides and polypeptides comprising epitope-bearing portions of the polypeptides of the present invention. These epitopes are immunogenic or antigenic epitopes of the polypeptides of the present invention. An "immunogenic epitope" is defined as a part of a protein that elicits an antibody response *in vivo* when the whole polypeptide of the present invention, or fragment thereof, is the immunogen. On the other hand, a region of a polypeptide to which an antibody can bind is defined as an "antigenic determinant" or "antigenic epitope." The number of *in vivo* immunogenic epitopes of a protein generally is less than the number of antigenic epitopes. See, e.g., Geysen, et al. (1983) Proc. Natl. Acad. Sci. USA 81:3998- 4002. However, antibodies can be made to any antigenic epitope, regardless of whether it is an immunogenic epitope, by using methods such as phage display. See e.g., Petersen G. et al. (1995) Mol. Gen. Genet. 249:425-431. Therefore, included in the present invention are both immunogenic epitopes and antigenic epitopes.

A list of exemplified amino acid sequences comprising immunogenic epitopes are shown in Table 1 below. It is pointed out that Table 1 only lists amino acid residues comprising epitopes predicted to have the highest degree of antigenicity using the algorithm of Jameson and Wolf, (1988) Comp. Appl. Biosci. 4:181-186 (said

references incorporated by reference in their entireties). The Jameson-Wolf antigenic analysis was performed using the computer program PROTEAN, using default parameters (Version 3.11 for the Power Macintosh, DNASTAR, Inc., 1228 South Park Street Madison, WI). Table 1 and portions of polypeptides not listed in Table 1 are not considered non-immunogenic. The immunogenic epitopes of Table 1 is an exemplified list, not an exhaustive list, because other immunogenic epitopes are merely not recognized as such by the particular algorithm used. Amino acid residues comprising other immunogenic epitopes may be routinely determined using algorithms similar to the Jameson-Wolf analysis or by *in vivo* testing for an antigenic response using methods known in the art. See, e.g., Geysen et al., *supra*; U.S. Patents 4,708,781; 5,194,392; 4,433,092; and 5,480,971 (said references incorporated by reference in their entireties). As shown in Table 1, SEQ ID NO:2 was found antigenic at amino acids: 22-29; 48-56; 62-73; 78-85; 88-95; 99-105; 118-126; 139-146; 151-169; 188-206; 208-231; 264-271; 286-293; 300-313; 317-342; 347-353; 363-369.

It is particularly pointed out that the amino acid sequences of Table 1 comprise immunogenic epitopes. Table 1 lists only the critical residues of immunogenic epitopes determined by the Jameson-Wolf analysis. Thus, additional flanking residues on either the N-terminal, C-terminal, or both N- and C-terminal ends may be added to the sequences of Table 1 to generate an epitope-bearing polypeptide of the present invention. Therefore, the immunogenic epitopes of Table 1 may include additional N-terminal or C-terminal amino acid residues. The additional flanking amino acid residues may be contiguous flanking N-terminal and/or C-terminal sequences from the polypeptides of the present invention, heterologous polypeptide sequences, or may include both contiguous flanking sequences from the polypeptides of the present invention and heterologous polypeptide sequences. Polypeptides of the present invention comprising immunogenic or antigenic epitopes are at least 7 amino acids residues in length. "At least" means that a polypeptide of the present invention comprising an immunogenic or antigenic epitope may be 7 amino acid residues in length or any integer between 7 amino acids and the number of amino acid residues of the full length polypeptides of the invention. Preferred polypeptides comprising immunogenic or antigenic epitopes are at least 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100 amino acid residues in length. However, it is pointed out that each and every integer between 7 and the number of amino acid residues of the full length polypeptide are included in the present invention.

The immuno and antigenic epitope-bearing fragments may be specified by either the number of contiguous amino acid residues, as described above, or further

specified by N-terminal and C-terminal positions of these fragments on the amino acid sequence of SEQ ID NO:2. Every combination of a N-terminal and C-terminal position that a fragment of, for example, at least 7 or at least 15 contiguous amino acid residues in length could occupy on the amino acid sequence of SEQ ID NO:2 is included in the invention. Again, "at least 7 contiguous amino acid residues in length" means 7 amino acid residues in length or any integer between 7 amino acids and the number of amino acid residues of the full length polypeptide of the present invention. Specifically, each and every integer between 7 and the number of amino acid residues of the full length polypeptide are included in the present invention.

Immunogenic and antigenic epitope-bearing polypeptides of the invention are useful, for example, to make antibodies which specifically bind the polypeptides of the invention, and in immunoassays to detect the polypeptides of the present invention. The antibodies are useful, for example, in affinity purification of the polypeptides of the present invention. The antibodies may also routinely be used in a variety of qualitative or quantitative immunoassays, specifically for the polypeptides of the present invention using methods known in the art. *See, e.g.,* Harlow et al., *Antibodies: A Laboratory Manual*, (Cold Spring Harbor Laboratory Press; 2nd Ed. 1988).

The epitope-bearing polypeptides of the present invention may be produced by any conventional means for making polypeptides including synthetic and recombinant methods known in the art. For instance, epitope-bearing peptides may be synthesized using known methods of chemical synthesis. For instance, Houghten has described a simple method for the synthesis of large numbers of peptides, such as 10-20 mgs of 248 individual and distinct 13 residue peptides representing single amino acid variants of a segment of the HA1 polypeptide, all of which were prepared and characterized (by ELISA-type binding studies) in less than four weeks (Houghten, R. A. *Proc. Natl. Acad. Sci. USA* 82:5131-5135 (1985)). This "Simultaneous Multiple Peptide Synthesis (SMPS)" process is further described in U.S. Patent No. 4,631,211 to Houghten and coworkers (1986). In this procedure the individual resins for the solid-phase synthesis of various peptides are contained in separate solvent-permeable packets, enabling the optimal use of the many identical repetitive steps involved in solid-phase methods. A completely manual procedure allows 500-1000 or more syntheses to be conducted simultaneously (Houghten et al. (1985) *Proc. Natl. Acad. Sci.* 82:5131-5135 at 5134).

Epitope-bearing polypeptides of the present invention are used to induce antibodies according to methods well known in the art including, but not limited to, *in vivo* immunization, *in vitro* immunization, and phage display methods. *See, e.g.,*

Suteliffe, et al., *supra*; Wilson, et al., *supra*, and Bittle, et al. (1985) J. Gen. Virol. 66:2347-2354. If *in vivo* immunization is used, animals may be immunized with free peptide; however, anti-peptide antibody titer may be boosted by coupling of the peptide to a macromolecular carrier, such as keyhole limpet hemacyanin (KLH) or tetanus toxoid. For instance, peptides containing cysteine residues may be coupled to a carrier using a linker such as -maleimidobenzoyl- N-hydroxysuccinimide ester (MBS), while other peptides may be coupled to carriers using a more general linking agent such as glutaraldehyde. Animals such as rabbits, rats and mice are immunized with either free or carrier-coupled peptides, for instance, by intraperitoneal and/or intradermal injection of emulsions containing about 100 µgs of peptide or carrier protein and Freund's adjuvant. Several booster injections may be needed, for instance, at intervals of about two weeks, to provide a useful titer of anti-peptide antibody which can be detected, for example, by ELISA assay using free peptide adsorbed to a solid surface. The titer of anti-peptide antibodies in serum from an immunized animal may be increased by selection of anti-peptide antibodies, for instance, by adsorption to the peptide on a solid support and elution of the selected antibodies according to methods well known in the art.

As one of skill in the art will appreciate, and discussed above, the polypeptides of the present invention comprising an immunogenic or antigenic epitope can be fused to heterologous polypeptide sequences. For example, the polypeptides of the present invention may be fused with the constant domain of immunoglobulins (IgA, IgE, IgG, IgM), or portions thereof (CH1, CH2, CH3, any combination thereof including both entire domains and portions thereof) resulting in chimeric polypeptides. These fusion proteins facilitate purification, and show an increased half-life *in vivo*. This has been shown, *e.g.*, for chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. *See, e.g.*, EPA 0,394,827; Traunecker et al. (1988) Nature 331:84-86. Fusion proteins that have a disulfide-linked dimeric structure due to the IgG portion can also be more efficient in binding and neutralizing other molecules than monomeric polypeptides or fragments thereof alone. *See, e.g.*, Fountoulakis et al. (1995) J. Biochem. 270:3958-3964. Nucleic acids encoding the above epitopes can also be recombined with a gene of interest as an epitope tag to aid in detection and purification of the expressed polypeptide.

### 35 Antibodies

The present invention further relates to antibodies and T-cell antigen receptors (TCR) which specifically bind the polypeptides of the present invention. The



antibodies of the present invention include IgG (including IgG1, IgG2, IgG3, and IgG4), IgA (including IgA1 and IgA2), IgD, IgE, or IgM, and IgY. As used herein, the term "antibody" (Ab) is meant to include whole antibodies, including single-chain whole antibodies, and antigen-binding fragments thereof. Most preferably the

5 antibodies are human antigen binding antibody fragments of the present invention include, but are not limited to, Fab, Fab' and F(ab')<sub>2</sub>, Fd, single-chain Fvs (scFv), single-chain antibodies, disulfide-linked Fvs (sdFv) and fragments comprising either a V<sub>L</sub> or V<sub>H</sub> domain. The antibodies may be from any animal origin including birds and mammals. Preferably, the antibodies are human, murine, rabbit, goat, guinea pig,

10 camel, horse, or chicken.

Antigen-binding antibody fragments, including single-chain antibodies, may comprise the variable region(s) alone or in combination with the entire or partial of the following: hinge region, CH1, CH2, and CH3 domains. Also included in the invention are any combinations of variable region(s) and hinge region, CH1, CH2,

15 and CH3 domains. The present invention further includes chimeric, humanized, and human monoclonal and polyclonal antibodies which specifically bind the polypeptides of the present invention. The present invention further includes antibodies which are anti-idiotypic to the antibodies of the present invention.

The antibodies of the present invention may be monospecific, bispecific,

20 trispecific or of greater multispecificity. Multispecific antibodies may be specific for different epitopes of a polypeptide of the present invention or may be specific for both a polypeptide of the present invention as well as for heterologous compositions, such as a heterologous polypeptide or solid support material. *See, e.g.*, WO 93/17715; WO 92/08802; WO 91/00360; WO 92/05793; Tutt, A. et al. (1991) J. Immunol. 147:60-

25 69; US Patents 5,573,920, 4,474,893, 5,601,819, 4,714,681, 4,925,648; Kostelny, S.A. et al. (1992) J. Immunol. 148:1547-1553.

Antibodies of the present invention may be described or specified in terms of the epitope(s) or portion(s) of a polypeptide of the present invention which are recognized or specifically bound by the antibody. The epitope(s) or polypeptide

30 portion(s) may be specified as described herein, *e.g.*, by N-terminal and C-terminal positions, by size in contiguous amino acid residues, or listed in the Tables and Figures. Antibodies which specifically bind any epitope or polypeptide of the present invention may also be excluded. Therefore, the present invention includes antibodies that specifically bind polypeptides of the present invention, and allows for the

35 exclusion of the same.

Antibodies of the present invention may also be described or specified in terms of their cross-reactivity. Antibodies that do not bind any other analog, ortholog, or

homolog of the polypeptides of the present invention are included. Antibodies that do not bind polypeptides with less than 95%, less than 90%, less than 85%, less than 80%, less than 75%, less than 70%, less than 65%, less than 60%, less than 55%, and less than 50% identity (as calculated using methods known in the art and described herein) to a polypeptide of the present invention are also included in the present invention. Further included in the present invention are antibodies which only bind polypeptides encoded by polynucleotides which hybridize to a polynucleotide of the present invention under stringent hybridization conditions (as described herein). Antibodies of the present invention may also be described or specified in terms of their binding affinity. Preferred binding affinities include those with a dissociation constant or  $K_d$  less than  $5 \times 10^{-6} M$ ,  $10^{-6} M$ ,  $5 \times 10^{-7} M$ ,  $10^{-7} M$ ,  $5 \times 10^{-8} M$ ,  $10^{-8} M$ ,  $5 \times 10^{-9} M$ ,  $10^{-9} M$ ,  $5 \times 10^{-10} M$ ,  $10^{-10} M$ ,  $5 \times 10^{-11} M$ ,  $10^{-11} M$ ,  $5 \times 10^{-12} M$ ,  $10^{-12} M$ ,  $5 \times 10^{-13} M$ ,  $10^{-13} M$ ,  $5 \times 10^{-14} M$ ,  $10^{-14} M$ ,  $5 \times 10^{-15} M$ , and  $10^{-15} M$ .

Antibodies of the present invention have uses that include, but are not limited to, methods known in the art to purify, detect, and target the polypeptides of the present invention including both *in vitro* and *in vivo* diagnostic and therapeutic methods. For example, the antibodies have use in immunoassays for qualitatively and quantitatively measuring levels of the polypeptides of the present invention in biological samples. See, e.g., Harlow et al., *ANTIBODIES: A LABORATORY MANUAL*, (Cold Spring Harbor Laboratory Press, 2nd ed. 1988) (incorporated by reference in the entirety).

The antibodies of the present invention may be used either alone or in combination with other compositions. The antibodies may further be recombinantly fused to a heterologous polypeptide at the N- or C-terminus or chemically conjugated (including covalently and non-covalently conjugations) to polypeptides or other compositions. For example, antibodies of the present invention may be recombinantly fused or conjugated to molecules useful as labels in detection assays and effector molecules such as heterologous polypeptides, drugs, or toxins. See, e.g., WO 92/08495; WO 91/14438; WO 89/12624; US Patent 5,314,995; and EP 0 396 387.

The antibodies of the present invention may be prepared by any suitable method known in the art. For example, a polypeptide of the present invention or an antigenic fragment thereof can be administered to an animal in order to induce the production of sera containing polyclonal antibodies. Monoclonal antibodies can be prepared using a wide of techniques known in the art including the use of hybridoma and recombinant technology. See, e.g., Harlow et al., *ANTIBODIES: A LABORATORY MANUAL*, (Cold Spring Harbor Laboratory Press, 2nd ed. 1988); Hammerling, et al., in: *MONOCLONAL ANTIBODIES AND T-CELL*

HYBRIDOMAS 563-681 (Elsevier, N.Y., 1981) (said references incorporated by reference in their entireties). Fab and Fab'(ab')<sub>2</sub> fragments may be produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce Fab'(ab')<sub>2</sub> fragments).

5 Alternatively, antibodies of the present invention can be produced through the application of recombinant DNA technology or through synthetic chemistry using methods known in the art. For example, the antibodies of the present invention can be prepared using various phage display methods known in the art. In phage display methods, functional antibody domains are displayed on the surface of a phage particle  
10 which carries polynucleotide sequences encoding them. Phage with a desired binding property are selected from a repertoire or combinatorial antibody library (e.g. human or murine) by selecting directly with antigen, typically antigen bound or captured to a solid surface or head. Phage used in these methods are typically filamentous phage including fd and M13 with Fab, Fv or disulfide stabilized Fv antibody domains  
15 recombinantly fused to either the phage gene III or gene VIII protein. Examples of phage display methods that can be used to make the antibodies of the present invention include those disclosed in Brinkman U. et al. (1995) *J. Immunol. Methods* 182:41-50; Ames, R.S. et al. (1995) *J. Immunol. Methods* 184:177-186; Kettleborough, C.A. et al. (1994) *Eur. J. Immunol.* 24:952-958; Persic, L. et al. (1997) *Gene* 187 9-18; Burton, D.R. et al. (1994) *Advances in Immunology* 57:191-280;  
20 PCT/GB91/01134; WO 90/02809; WO 91/10737; WO 92/01047; WO 92/18619; WO 93/11236; WO 95/15982; WO 95/20401; and US Patents 5,698,426, 5,223,409, 5,403,484, 5,580,717, 5,427,908, 5,750,753, 5,821,047, 5,571,698, 5,427,908, 5,516,637, 5,780,225, 5,658,727 and 5,733,743 (said references incorporated by  
25 reference in their entireties).

As described in the above references, after phage selection, the antibody coding regions from the phage can be isolated and used to generate whole antibodies, including human antibodies, or any other desired antigen binding fragment, and expressed in any desired host including mammalian cells, insect cells, plant cells,  
30 yeast, and bacteria. For example, techniques to recombinantly produce Fab, Fab' and F(ab')<sub>2</sub> fragments can also be employed using methods known in the art such as those disclosed in WO 92/22324; Mullinax, R.L. et al. (1992) *BioTechniques* 12(6):864-869; and Sawai, H. et al. (1995) *AJRI* 34:26-34; and Better, M. et al. (1988) *Science* 240:1041-1043 (said references incorporated by reference in their entireties).

35 Examples of techniques which can be used to produce single-chain Fvs and antibodies include those described in U.S. Patents 4,946,778 and 5,258,498; Huston et al. (1991) *Methods in Enzymology* 203:46-88; Shu, L. et al. (1993) *PNAS*

90:7995-7999; and Skerra, A. et al. (1988) Science 240:1038-1040. For some uses, including *in vivo* use of antibodies in humans and *in vitro* detection assays, it may be preferable to use chimeric, humanized, or human antibodies. Methods for producing chimeric antibodies are known in the art. See e.g., Morrison, Science 229:1202  
5 (1985); Oi et al., BioTechniques 4:214 (1986); Gillies, S.D. et al. (1989) J. Immunol. Methods 125:191-202; and US Patent 5,807,715. Antibodies can be humanized using a variety of techniques including CDR grafting (EP 0 239 400; WO 91/09967, US Patent 5,530,101; and 5,585,089), veneering or resurfacing (EP 0 592 106; EP 0 519 596; Padlan E.A., (1991) Molecular Immunology 28(4/5):489-498; Studnicka  
10 G.M. et al. (1994) Protein Engineering 7(6):805-814; Roguska M.A. et al. (1994) PNAS 91:969-973), and chain shuffling (US Patent 5,565,332). Human antibodies can be made by a variety of methods known in the art including phage display methods described above. See also, US Patents 4,444,887, 4,716,111, 5,545,806, and 5,814,318; and WO 98/46645 (said references incorporated by reference in their  
15 entireties).

Further included in the present invention are antibodies recombinantly fused or chemically conjugated (including both covalently and non-covalently conjugations) to a polypeptide of the present invention. The antibodies may be specific for antigens other than polypeptides of the present invention. For example, antibodies may be  
20 used to target the polypeptides of the present invention to particular cell types, either *in vitro* or *in vivo*, by fusing or conjugating the polypeptides of the present invention to antibodies specific for particular cell surface receptors. Antibodies fused or conjugated to the polypeptides of the present invention may also be used in *in vitro* immunoassays and purification methods using methods known in the art. See e.g.,  
25 Harbor et al. *supra* and WO 93/21232; EP 0 439 095; Naramura, M. et al. (1994) Immunol. Lett. 39:91-99; US Patent 5,474,981; Gillies, S.O. et al. (1992) PNAS 89:1428-1432; Fell, H.P. et al. (1991) J. Immunol. 146:2446-2452 (said references incorporated by reference in their entireties).

The present invention further includes compositions comprising the  
30 polypeptides of the present invention fused or conjugated to antibody domains other than the variable regions. For example, the polypeptides of the present invention may be fused or conjugated to an antibody Fc region, or portion thereof. The antibody portion fused to a polypeptide of the present invention may comprise the hinge region, CH1 domain, CH2 domain, and CH3 domain or any combination of whole domains  
35 or portions thereof. The polypeptides of the present invention may be fused or conjugated to the above antibody portions to increase the *in vivo* half life of the polypeptides or for use in immunoassays using methods known in the art. The

polypeptides may also be fused or conjugated to the above antibody portions to form multimers. For example, Fc portions fused to the polypeptides of the present invention can form dimers through disulfide bonding between the Fc portions. Higher multimeric forms can be made by fusing the polypeptides to portions of IgA and IgM. Methods for fusing or conjugating the polypeptides of the present invention to antibody portions are known in the art. *See e.g.*, US Patents 5,336,603, 5,622,929, 5,359,046, 5,349,053, 5,447,851, 5,112,946; EP 0 307 434, EP 0 367 166; WO 96/04388, WO 91/06570; Ashkenazi, A. et al. (1991) PNAS 88:10535-10539; Zheng, X.X. et al. (1995) J. Immunol. 154:5590-5600; and Vil, H. et al. (1992) PNAS 89:11337-11341 (said references incorporated by reference in their entireties).

The invention further relates to antibodies which act as agonists or antagonists of the polypeptides of the present invention. For example, the present invention includes antibodies which disrupt the receptor/ligand interactions with the polypeptides of the invention either partially or fully. Included are both receptor-specific antibodies and ligand-specific antibodies. Included are receptor-specific antibodies which do not prevent ligand binding but prevent receptor activation. Receptor activation (i.e., signaling) may be determined by techniques described herein or otherwise known in the art. Also include are receptor-specific antibodies which both prevent ligand binding and receptor activation. Likewise, included are neutralizing antibodies which bind the ligand and prevent binding of the ligand to the receptor, as well as antibodies which bind the ligand, thereby preventing receptor activation, but do not prevent the ligand from binding the receptor. Further included are antibodies which activate the receptor. These antibodies may act as agonists for either all or less than all of the biological activities affected by ligand-mediated receptor activation. The antibodies may be specified as agonists or antagonists for biological activities comprising specific activities disclosed herein. The above antibody agonists can be made using methods known in the art. *See e.g.*, WO 96/40281; US Patent 5,811,097; Deng, B. et al. (1998) Blood 92(6):1981-1988; Chen, Z. et al. (1998) Cancer Res. 58(16):3668-3678; Harrop, J.A. et al. (1998) J. Immunol. 161(4):1786-1794; Zhu, Z. et al. (1998) Cancer Res. 58(15):3209-3214; Yoon, D.Y. et al. (1998) J. Immunol. 160(7):3170-3179; Prat, M. et al. (1998) J. Cell. Sci. 111(Pt2):237-247; Pitard, V. et al. (1997) J. Immunol. Methods 205(2):177-190; Liautard, J. et al. (1997) Cytokine 9(4):233-241; Carlson, N.G. et al. (1997) J. Biol. Chem. 272(17):11295-11301; Taryman, R.E. et al. (1995) Neuron 14(4):755-762; Muller, Y.A. et al. (1998) Structure 6(9):1153-1167; Bartunek, P. et al. (1996) Cytokine 8(1):14-20 (said references incorporated by reference in their entireties).

### Fusion Proteins

Any CRCGCL polypeptide can be used to generate fusion proteins. For example, the CRCGCL polypeptide, when fused to a second protein, can be used as an antigenic tag. Antibodies raised against the CRCGCL polypeptide can be used to indirectly detect the second protein by binding to the CRCGCL. Moreover, because secreted proteins target cellular locations based on trafficking signals, the CRCGCL polypeptides can be used as a targeting molecule once fused to other proteins.

Examples of domains that can be fused to CRCGCL polypeptides include not only heterologous signal sequences, but also other heterologous functional regions. The fusion does not necessarily need to be direct, but may occur through linker sequences.

Moreover, fusion proteins may also be engineered to improve characteristics of the CRCGCL polypeptide. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the CRCGCL polypeptide to improve stability and persistence during purification from the host cell or subsequent handling and storage. Also, peptide moieties may be added to the CRCGCL polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the CRCGCL polypeptide. The addition of peptide moieties to facilitate handling of polypeptides are familiar and routine techniques in the art.

Moreover, CRCGCL polypeptides, including fragments, and specifically epitopes, can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP A 394,827; Traunecker et al., Nature 331:84-86 (1988).) Fusion proteins having disulfide-linked dimeric structures (due to the IgG) can also be more efficient in binding and neutralizing other molecules, than the monomeric secreted protein or protein fragment alone. (Fountoulakis et al., J. Biochem. 270:3958-3964 (1995).)

Similarly, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP-A 0232 262.) Alternatively, deleting the Fc part after the fusion protein has been expressed, detected, and purified.

would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See, D. Bennett et al., *J. Molecular Recognition* 8:52-58 (1995); K. Johanson et al., *J. Biol. Chem.* 270:9459-9471 (1995).)

Moreover, the CRCGCL polypeptides can be fused to marker sequences, such as a peptide which facilitates purification of CRCGCL. In preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA. 91311), among others, many of which are commercially available. As described in Gentz et al., *Proc. Natl. Acad. Sci. USA* 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. Another peptide tag useful for purification, the "HA" tag, corresponds to an epitope derived from the influenza hemagglutinin protein. (Wilson et al., *Cell* 37:767 (1984).)

Thus, any of these above fusions can be engineered using the CRCGCL polynucleotides or the polypeptides.

### **Vectors, Host Cells, and Protein Production**

The present invention also relates to vectors containing the CRCGCL polynucleotide, host cells, and the production of polypeptides by recombinant techniques. The vector may be, for example, a phage, plasmid, viral, or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

CRCGCL polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

The CRCGCL polynucleotide insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the *E. coli* lac, trp, *phoA* and *tac* promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably

include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance genes for culturing in *E. coli* and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as *E. coli*, *Streptomyces* and *Salmonella typhimurium* cells; fungal cells, such as yeast cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS, 293, and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from QIAGEN, Inc.; pBluescript vectors, Phagescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene Cloning Systems, Inc.; and ptre99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia Biotech, Inc. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL, available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., *Basic Methods In Molecular Biology* (1986). It is specifically contemplated that CRGCL polypeptides may in fact be expressed by a host cell lacking a recombinant vector.

CRGCL polypeptides can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

CRGCL polypeptides, and preferably the secreted form, can also be recovered from: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect, and mammalian cells. Depending upon the host employed in a recombinant production



procedure, the CRCGCL polypeptides may be glycosylated or may be non-glycosylated. In addition, CRCGCL polypeptides may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

In addition to encompassing host cells containing the vector constructs discussed herein, the invention also encompasses primary, secondary, and immortalized host cells of vertebrate origin, particularly mammalian origin, that have been engineered to delete or replace endogenous genetic material (e.g., CRCGCL coding sequence), and/or to include genetic material (e.g., heterologous polynucleotide sequences) that is operably associated with CRCGCL polynucleotides of the invention, and which activates, alters, and/or amplifies endogenous CRCGCL polynucleotides. For example, techniques known in the art may be used to operably associate heterologous control regions (e.g., promoter and/or enhancer) and endogenous CRCGCL polynucleotide sequences via homologous recombination (see, e.g., U.S. Patent No. 5,641,670, issued June 24, 1997; International Publication No. WO 96/29411, published September 26, 1996; International Publication No. WO 94/12650, published August 4, 1994; Kollier et al., Proc. Natl. Acad. Sci. USA 86:8932-8935 (1989); and Zijlstra et al., Nature 342:435-438 (1989), the disclosures of each of which are incorporated by reference in their entireties).

In addition, polypeptides of the invention can be chemically synthesized using techniques known in the art (e.g., see Creighton, 1983, Proteins: Structures and Molecular Principles, W.H. Freeman & Co., N.Y., and Hunkapiller, M., et al., 1984, Nature 310:105-111). For example, a peptide corresponding to a fragment of the CRCGCL polypeptides of the invention can be synthesized by use of a peptide synthesizer. Furthermore, if desired, nonclassical amino acids or chemical amino acid analogs can be introduced as a substitution or addition into the CRCGCL polynucleotide sequence. Non-classical amino acids include, but are not limited to, to the D-isomers of the common amino acids, 2,4-diaminobutyric acid,  $\alpha$ -amino isobutyric acid, 4-aminobutyric acid, Abu, 2-amino butyric acid,  $\gamma$ -Abu,  $\epsilon$ -Ahx, 6-amino hexanoic acid, Aib, 2-amino isobutyric acid, 3-amino propionic acid, ornithine, norleucine, norvaline, hydroxyproline, sarcosine, citrulline, homocitrulline, cysteic acid, t-butylglycine, t-butylalanine, phenylglycine, cyclohexylalanine, b-

alanine, fluoro-amino acids, designer amino acids such as  $\beta$ -methyl amino acids,  $\alpha$ -methyl amino acids,  $N$ -methyl amino acids, and amino acid analogs in general. Furthermore, the amino acid can be D (dextrorotary) or L (levorotary).

The invention encompasses CRGCL polypeptides which are differentially  
5 modified during or after translation, e.g., by glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to an antibody molecule or other cellular ligand, etc. Any of numerous chemical modifications may be carried out by known techniques, including but not limited, to specific chemical cleavage by cyanogen bromide, trypsin,  
10 chymotrypsin, papain, V8 protease,  $\text{NaBH}_4$ ; acetylation, formylation, oxidation, reduction; metabolic synthesis in the presence of tunicamycin; etc.

Additional post-translational modifications encompassed by the invention include, for example, e.g., N-linked or O-linked carbohydrate chains, processing of N-terminal or C-terminal ends, attachment of chemical moieties to the amino acid  
15 backbone, chemical modifications of N-linked or O-linked carbohydrate chains, and addition or deletion of an N-terminal methionine residue as a result of procaryotic host cell expression. The polypeptides may also be modified with a detectable label, such as an enzymatic, fluorescent, isotopic or affinity label to allow for detection and isolation of the protein.

Also provided by the invention are chemically modified derivatives of  
20 CRGCL which may provide additional advantages such as increased solubility, stability and circulating time of the polypeptide, or decreased immunogenicity (see U. S. Patent No. 4,179,337). The chemical moieties for derivitization may be selected from water soluble polymers such as polyethylene glycol, ethylene glycol/propylene  
25 glycol copolymers, carboxymethylcellulose, dextran, polyvinyl alcohol and the like. The polypeptides may be modified at random positions within the molecule, or at predetermined positions within the molecule and may include one, two, three or more attached chemical moieties.

The polymer may be of any molecular weight, and may be branched or  
30 unbranched. For polyethylene glycol, the preferred molecular weight is between about 1 kDa and about 100 kDa (the term "about" indicating that in preparations of polyethylene glycol, some molecules will weigh more, some less, than the stated molecular weight) for ease in handling and manufacturing. Other sizes may be used, depending on the desired therapeutic profile (e.g., the duration of sustained release  
35 desired, the effects, if any on biological activity, the ease in handling, the degree or lack of antigenicity and other known effects of the polyethylene glycol to a therapeutic protein or analog).

The polyethylene glycol molecules (or other chemical moieties) should be attached to the protein with consideration of effects on functional or antigenic domains of the protein. There are a number of attachment methods available to those skilled in the art, e.g., EP 0 401 384, herein incorporated by reference (coupling PEG to G-CSF), see also Malik et al., *Exp. Hematol.* 20:1028-1035 (1992) (reporting pegylation of GM-CSF using tresyl chloride). For example, polyethylene glycol may be covalently bound through amino acid residues via a reactive group, such as, a free amino or carboxyl group. Reactive groups are those to which an activated polyethylene glycol molecule may be bound. The amino acid residues having a free amino group may include lysine residues and the N-terminal amino acid residues; those having a free carboxyl group may include aspartic acid residues, glutamic acid residues and the C-terminal amino acid residue. Sulfhydryl groups may also be used as a reactive group for attaching the polyethylene glycol molecules. Preferred for therapeutic purposes is attachment at an amino group, such as attachment at the N-terminus or lysine group.

One may specifically desire proteins chemically modified at the N-terminus. Using polyethylene glycol as an illustration of the present composition, one may select from a variety of polyethylene glycol molecules (by molecular weight, branching, etc.), the proportion of polyethylene glycol molecules to protein (or peptide) molecules in the reaction mix, the type of pegylation reaction to be performed, and the method of obtaining the selected N-terminally pegylated protein. The method of obtaining the N-terminally pegylated preparation (i.e., separating this moiety from other monopegylated moieties if necessary) may be by purification of the N-terminally pegylated material from a population of pegylated protein molecules. Selective proteins chemically modified at the N-terminus modification may be accomplished by reductive alkylation which exploits differential reactivity of different types of primary amino groups (lysine versus the N-terminal) available for derivatization in a particular protein. Under the appropriate reaction conditions, substantially selective derivatization of the protein at the N-terminus with a carbonyl group containing polymer is achieved.

The CRCGCL polypeptides of the invention may be in monomers or multimers (i.e., dimers, trimers, tetramers and higher multimers). Accordingly, the present invention relates to monomers and multimers of the CRCGCL polypeptides of the invention, their preparation, and compositions (preferably, pharmaceutical compositions) containing them. In specific embodiments, the polypeptides of the invention are monomers, dimers, trimers or tetramers. In additional embodiments, the multimers of the invention are at least dimers, at least trimers, or at least tetramers.

Multimers encompassed by the invention may be homomers or heteromers. As used herein, the term homomer, refers to a multimer containing only CRCGCL polypeptides of the invention (including CRCGCL fragments, variants, splice variants, and fusion proteins, as described herein). These homomers may contain  
5 CRCGCL polypeptides having identical or different amino acid sequences. In a specific embodiment, a homomer of the invention is a multimer containing only CRCGCL polypeptides having an identical amino acid sequence. In another specific embodiment, a homomer of the invention is a multimer containing CRCGCL polypeptides having different amino acid sequences. In specific embodiments, the  
10 multimer of the invention is a homodimer (e.g., containing CRCGCL polypeptides having identical or different amino acid sequences) or a homotrimer (e.g., containing CRCGCL polypeptides having identical and/or different amino acid sequences). In additional embodiments, the homomeric multimer of the invention is at least a homodimer, at least a homotrimer, or at least a homotetramer.

15 As used herein, the term heteromer refers to a multimer containing one or more heterologous polypeptides (i.e., polypeptides of different proteins) in addition to the CRCGCL polypeptides of the invention. In a specific embodiment, the multimer of the invention is a heterodimer, a heterotrimer, or a heterotetramer. In additional embodiments, the homomeric multimer of the invention is at least a homodimer, at  
20 least a homotrimer, or at least a homotetramer.

Multimers of the invention may be the result of hydrophobic, hydrophilic, ionic and/or covalent associations and/or may be indirectly linked, by for example, liposome formation. Thus, in one embodiment, multimers of the invention, such as, for example, homodimers or homotrimers, are formed when polypeptides of the  
25 invention contact one another in solution. In another embodiment, heteromultimers of the invention, such as, for example, heterotrimers or heterotetramers, are formed when polypeptides of the invention contact antibodies to the polypeptides of the invention (including antibodies to the heterologous polypeptide sequence in a fusion protein of the invention) in solution. In other embodiments, multimers of the  
30 invention are formed by covalent associations with and/or between the CRCGCL polypeptides of the invention. Such covalent associations may involve one or more amino acid residues contained in the polypeptide sequence (e.g., that recited in SEQ ID NO:2, or contained in the polypeptide encoded by the clone HTAEK53). In one instance, the covalent associations are cross-linking between cysteine residues located  
35 within the polypeptide sequences which interact in the native (i.e., naturally occurring) polypeptide. In another instance, the covalent associations are the consequence of chemical or recombinant manipulation. Alternatively, such covalent associations may

involve one or more amino acid residues contained in the heterologous polypeptide sequence in a CRCGCL fusion protein. In one example, covalent associations are between the heterologous sequence contained in a fusion protein of the invention (see, e.g., US Patent Number 5,478,925). In a specific example, the covalent associations are between the heterologous sequence contained in a CRCGCL-Fe fusion protein of the invention (as described herein). In another specific example, covalent associations of fusion proteins of the invention are between heterologous polypeptide sequence from another TNF family ligand/receptor member that is capable of forming covalently associated multimers, such as for example, osteoprotegerin (see, e.g., International Publication No. WO 98/49305, the contents of which are herein incorporated by reference in its entirety).

The multimers of the invention may be generated using chemical techniques known in the art. For example, polypeptides desired to be contained in the multimers of the invention may be chemically cross-linked using linker molecules and linker molecule length optimization techniques known in the art (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety). Additionally, multimers of the invention may be generated using techniques known in the art to form one or more inter-molecule cross-links between the cysteine residues located within the sequence of the polypeptides desired to be contained in the multimer (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety). Further, polypeptides of the invention may be routinely modified by the addition of cysteine or biotin to the C terminus or N-terminus of the polypeptide and techniques known in the art may be applied to generate multimers containing one or more of these modified polypeptides (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety). Additionally, techniques known in the art may be applied to generate liposomes containing the polypeptide components desired to be contained in the multimer of the invention (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety).

Alternatively, multimers of the invention may be generated using genetic engineering techniques known in the art. In one embodiment, polypeptides contained in multimers of the invention are produced recombinantly using fusion protein technology described herein or otherwise known in the art (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety). In a specific embodiment, polynucleotides coding for a homodimer of the invention are generated by ligating a polynucleotide sequence encoding a polypeptide of the invention to a sequence encoding a linker polypeptide and then further to a synthetic polynucleotide encoding the translated product of the polypeptide in the reverse

orientation from the original C-terminus to the N-terminus (lacking the leader sequence) (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety). In another embodiment, recombinant techniques described herein or otherwise known in the art are applied to generate recombinant polypeptides of the invention which contain a transmembrane domain (or hydrophobic or signal peptide) and which can be incorporated by membrane reconstitution techniques into liposomes (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety).

#### 10 Uses of the CRCGCL Polynucleotides

The CRCGCL polynucleotides identified herein can be used in numerous ways as reagents. The following description should be considered exemplary and utilizes known techniques.

There exists an ongoing need to identify new chromosome markers, since few chromosome marking reagents, based on actual sequence data (repeat polymorphisms), are presently available. Using a panel of radiation hybrids, CRCGCL maps to the pseudoautosomal region (PAR) of the sex chromosomes, which is located on both X (Xp22.3) and Y (Yp13.3). Interestingly, two other cytokine receptors map to this region (IL3Ra, and GMCSFRa). See, Kremer et al. "A Cytokine Receptor Gene Cluster in the X-Y pseudoautosomal region?" Blood 82(1) 22-28 (1993). Thus, CRCGCL polynucleotides can be used in linkage analysis as a marker for the pseudoautosomal region on the X and Y chromosomes.

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the sequences shown in SEQ ID NO:1. Primers can be selected using computer analysis so that primers do not span more than one predicted exon in the genomic DNA. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human CRCGCL gene corresponding to the SEQ ID NO:1 will yield an amplified fragment.

Similarly, somatic hybrids provide a rapid method of PCR mapping the polynucleotides to particular chromosomes. Three or more clones can be assigned per day using a single thermal cycler. Moreover, sublocalization of the CRCGCL polynucleotides can be achieved with panels of specific chromosome fragments. Other gene mapping strategies that can be used include in situ hybridization, prescreening with labeled flow-sorted chromosomes, and preselection by hybridization to construct chromosome specific-cDNA libraries.

Precise chromosomal location of the CRCGCL polynucleotides can also be achieved using fluorescence in situ hybridization (FISH) of a metaphase chromosomal spread. This technique uses polynucleotides as short as 500 or 600 bases, however, polynucleotides 2,000-4,000 bp are preferred. For a review of this technique, see Verma et al., "Human Chromosomes: a Manual of Basic Techniques," Pergamon Press, New York (1988).

For chromosome mapping, the CRCGCL polynucleotides can be used individually (to mark a single chromosome or a single site on that chromosome) or in panels (for marking multiple sites and/or multiple chromosomes). Preferred polynucleotides correspond to the noncoding regions of the cDNAs because the coding sequences are more likely conserved within gene families, thus increasing the chance of cross hybridization during chromosomal mapping.

Once a polynucleotide has been mapped to a precise chromosomal location, the physical position of the polynucleotide can be used in linkage analysis. Linkage analysis establishes coinheritance between a chromosomal location and presentation of a particular disease. (Disease mapping data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available on line through Johns Hopkins University Welch Medical Library) .) Assuming 1 megabase mapping resolution and one gene per 20 kb, a cDNA precisely localized to a chromosomal region associated with the disease could be one of 50-500 potential causative genes.

Thus, once coinheritance is established, differences in the CRCGCL polynucleotide and the corresponding gene between affected and unaffected individuals can be examined. First, visible structural alterations in the chromosomes, such as deletions or translocations, are examined in chromosome spreads or by PCR. If no structural alterations exist, the presence of point mutations are ascertained. Mutations observed in some or all affected individuals, but not in normal individuals, indicates that the mutation may cause the disease. However, complete sequencing of the CRCGCL polypeptide and the corresponding gene from several normal individuals is required to distinguish the mutation from a polymorphism. If a new polymorphism is identified, this polymorphic polypeptide can be used for further linkage analysis.

Furthermore, increased or decreased expression of the gene in affected individuals as compared to unaffected individuals can be assessed using CRCGCL polynucleotides. Any of these alterations (altered expression, chromosomal rearrangement, or mutation) can be used as a diagnostic or prognostic marker.

In addition to the foregoing, a CRCGCL polynucleotide can be used to control gene expression through triple helix formation or antisense DNA or RNA. Both

methods rely on binding of the polynucleotide to DNA or RNA. For these techniques, preferred polynucleotides are usually 20 to 40 bases in length and complementary to either the region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991), Oligodeoxy-nucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988).) Triple helix formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques are effective in model systems, and the information disclosed herein can be used to design antisense or triple helix polynucleotides in an effort to treat disease.

CRCGCL polynucleotides are also useful in gene therapy. One goal of gene therapy is to insert a normal gene into an organism having a defective gene, in an effort to correct the genetic defect. CRCGCL offers a means of targeting such genetic defects in a highly accurate manner. Another goal is to insert a new gene that was not present in the host genome, thereby producing a new trait in the host cell.

The CRCGCL polynucleotides are also useful for identifying individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identifying personnel. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The CRCGCL polynucleotides can be used as additional DNA markers for RFLP.

The CRCGCL polynucleotides can also be used as an alternative to RFLP, by determining the actual base-by-base DNA sequence of selected portions of an individual's genome. These sequences can be used to prepare PCR primers for amplifying and isolating such selected DNA, which can then be sequenced. Using this technique, individuals can be identified because each individual will have a unique set of DNA sequences. Once an unique ID database is established for an individual, positive identification of that individual, living or dead, can be made from extremely small tissue samples.

Forensic biology also benefits from using DNA-based identification techniques as disclosed herein. DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, semen,



etc., can be amplified using PCR. In one prior art technique, gene sequences amplified from polymorphic loci, such as DQa class II HLA gene, are used in forensic biology to identify individuals. (Erich, II., PCR Technology, Freeman and Co. (1992).) Once these specific polymorphic loci are amplified, they are digested with one or more restriction enzymes, yielding an identifying set of bands on a Southern blot probed with DNA corresponding to the DQa class II HLA gene. Similarly, CRCGCL polynucleotides can be used as polymorphic markers for forensic purposes.

There is also a need for reagents capable of identifying the source of a particular tissue. Such need arises, for example, in forensics when presented with tissue of unknown origin. Appropriate reagents can comprise, for example, DNA probes or primers specific to particular tissue prepared from CRCGCL sequences. Panels of such reagents can identify tissue by species and/or by organ type. In a similar fashion, these reagents can be used to screen tissue cultures for contamination.

Because CRCGCL is found expressed in a cervical cancer cell line (HeLa), activated T cells, and a lung carcinoma cell line (A549), while a shorter variant is also expressed in the lymph node and to a lesser extent in the spleen, CRCGCL polynucleotides are useful as hybridization probes for differential identification of the tissue(s) or cell type(s) present in a biological sample. Similarly, polypeptides and antibodies directed to CRCGCL polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). In addition, for a number of disorders of the above tissues or cells, particularly of the immune system, significantly higher or lower levels of CRCGCL gene expression may be detected in certain tissues (e.g., cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to a "standard" CRCGCL gene expression level, i.e., the CRCGCL expression level in healthy tissue from an individual not having the immune system disorder.

Thus, the invention provides a diagnostic method of a disorder, which involves: (a) assaying CRCGCL gene expression level in cells or body fluid of an individual; (b) comparing the CRCGCL gene expression level with a standard CRCGCL gene expression level, whereby an increase or decrease in the assayed CRCGCL gene expression level compared to the standard expression level is indicative of disorder in the immune system.

In the very least, the CRCGCL polynucleotides can be used as molecular weight markers on Southern gels, as diagnostic probes for the presence of a specific mRNA in a particular cell type, as a probe to "subtract-out" known sequences in the

process of discovering novel polynucleotides, for selecting and making oligomers for attachment to a 'gene chip' or other support, to raise anti-DNA antibodies using DNA immunization techniques, and as an antigen to elicit an immune response.

## 5 Uses of CRCGCL Polypeptides

CRCGCL polypeptides can be used in numerous ways. The following description should be considered exemplary and utilizes known techniques.

CRCGCL polypeptides can be used to assay protein levels in a biological sample using antibody-based techniques. For example, protein expression in tissues  
10 can be studied with classical immunohistological methods. (Jalkanen, M., et al., J. Cell. Biol. 101:976-985 (1985), Jalkanen, M., et al., J. Cell. Biol. 105:3087-3096 (1987).) Other antibody based methods useful for detecting protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and  
15 include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine (<sup>125</sup>I, <sup>121</sup>I), carbon (<sup>14</sup>C), sulfur (<sup>35</sup>S), tritium (<sup>3</sup>H), indium (<sup>112</sup>In), and technetium (<sup>99m</sup>Tc), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

In addition to assaying secreted protein levels in a biological sample, proteins  
20 can also be detected in vivo by imaging. Antibody labels or markers for in vivo imaging of protein include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as  
25 deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma

A protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example, <sup>131</sup>I, <sup>112</sup>In, <sup>99m</sup>Tc), a radio-opaque substance, or a material detectable by nuclear  
30 magnetic resonance, is introduced (for example, parenterally, subcutaneously, or intraperitoneally) into the mammal. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5  
35 to 20 millicuries of <sup>99m</sup>Tc. The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein. In vivo tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics

of Radiolabeled Antibodies and Their Fragments." (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S.W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982).)

Thus, the invention provides a diagnostic method of a disorder, which involves (a) assaying the expression of CRCGCL polypeptide in cells or body fluid of an individual; (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed CRCGCL polypeptide gene expression level compared to the standard expression level is indicative of a disorder.

Moreover, CRCGCL polypeptides can be used to treat disease. For example, patients can be administered CRCGCL polypeptides in an effort to replace absent or decreased levels of the CRCGCL polypeptide (e.g., insulin), to supplement absent or decreased levels of a different polypeptide (e.g., hemoglobin S for hemoglobin B), to inhibit the activity of a polypeptide (e.g., an oncogene), to activate the activity of a polypeptide (e.g., by binding to a receptor), to reduce the activity of a membrane bound receptor by competing with it for free ligand (e.g., soluble TNF receptors used in reducing inflammation), or to bring about a desired response (e.g., blood vessel growth).

Similarly, antibodies directed to CRCGCL polypeptides can also be used to treat disease. For example, administration of an antibody directed to a CRCGCL polypeptide can bind and reduce overproduction of the polypeptide. Similarly, administration of an antibody can activate the polypeptide, such as by binding to a polypeptide bound to a membrane (receptor).

At the very least, the CRCGCL polypeptides can be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. CRCGCL polypeptides can also be used to raise antibodies, which in turn are used to measure protein expression from a recombinant cell, as a way of assessing transformation of the host cell. Moreover, CRCGCL polypeptides can be used to test the following biological activities.

#### **Gene Therapy Methods**

Another aspect of the present invention is to gene therapy methods for treating disorders, diseases and conditions. The gene therapy methods relate to the introduction of nucleic acid (DNA, RNA and antisense DNA or RNA) sequences into an animal to achieve expression of the CRCGCL polypeptide of the present invention. This method requires a polynucleotide that codes for a CRCGCL polypeptide operatively linked to a promoter and any other genetic elements necessary for the

expression of the polypeptide by the target tissue. Such gene therapy and delivery techniques are known in the art, see, for example, WO90/11092, which is herein incorporated by reference.

Thus, for example, cells from a patient may be engineered with a  
5 polynucleotide (DNA or RNA) comprising a promoter operably linked to a CRGCL polynucleotide *et. in o.*, with the engineered cells then being provided to a patient to be treated with the polypeptide. Such methods are well-known in the art. For example, see Belldgrun, A., *et al.*, *J. Natl. Cancer Inst.* 85: 207-216 (1993); Ferrantini, M. *et al.*, *Cancer Research* 53: 1107-1112 (1993); Ferrantini, M. *et al.*, *J. Immunology*  
10 153: 4604-4615 (1994); Kaido, T., *et al.*, *Int. J. Cancer* 60: 221-229 (1995); Ogura, H., *et al.*, *Cancer Research* 50: 5102-5106 (1990); Santodonato, L., *et al.*, *Human Gene Therapy* 7:1-10 (1996); Santodonato, L., *et al.*, *Gene Therapy* 4:1246-1255 (1997); and Zhang, J.-F., *et al.*, *Cancer Gene Therapy* 3: 31-38 (1996)), which are herein incorporated by reference. In one embodiment, the cells that are engineered are  
15 arterial cells. The arterial cells may be reintroduced into the patient through direct injection to the artery, the tissues surrounding the artery, or through catheter injection.

As discussed in more detail below, the CRGCL polynucleotide constructs can be delivered by any method that delivers injectable materials to the cells of an animal, such as, injection into the interstitial space of tissues (heart, muscle, skin,  
20 lung, liver, and the like). The CRGCL polynucleotide constructs may be delivered in a pharmaceutically acceptable liquid or aqueous carrier.

In one embodiment, the CRGCL polynucleotide is delivered as a naked polynucleotide. The term "naked" polynucleotide, DNA or RNA refers to sequences that are free from any delivery vehicle that acts to assist, promote or facilitate entry  
25 into the cell, including viral sequences, viral particles, liposome formulations, lipofectin or precipitating agents and the like. However, the CRGCL polynucleotides can also be delivered in liposome formulations and lipofectin formulations and the like can be prepared by methods well known to those skilled in the art. Such methods are described, for example, in U.S. Patent Nos. 5,593,972,  
30 5,589,466, and 5,580,859, which are herein incorporated by reference.

The CRGCL polynucleotide vector constructs used in the gene therapy method are preferably constructs that will not integrate into the host genome nor will they contain sequences that allow for replication. Appropriate vectors include pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; pSVK3,  
35 pBPV, pMSG and pSVL available from Pharmacia; and pEF1/V5, pcDNA3.1, and pRc/CMV2 available from Invitrogen. Other suitable vectors will be readily apparent to the skilled artisan.

Any strong promoter known to those skilled in the art can be used for driving the expression of CRGCL DNA. Suitable promoters include adenoviral promoters, such as the adenoviral major late promoter; or heterologous promoters, such as the cytomegalovirus (CMV) promoter; the respiratory syncytial virus (RSV) promoter; inducible promoters, such as the MMT promoter, the metallothionein promoter; heat shock promoters; the albumin promoter; the ApoA1 promoter; human globin promoters; viral thymidine kinase promoters, such as the Herpes Simplex thymidine kinase promoter; retroviral LTRs; the  $\beta$ -actin promoter; and human growth hormone promoters. The promoter also may be the native promoter for CRGCL.

Unlike other gene therapy techniques, one major advantage of introducing naked nucleic acid sequences into target cells is the transitory nature of the polynucleotide synthesis in the cells. Studies have shown that non replicating DNA sequences can be introduced into cells to provide production of the desired polypeptide for periods of up to six months.

The CRGCL polynucleotide construct can be delivered to the interstitial space of tissues within the an animal, including of muscle, skin, brain, lung, liver, spleen, bone marrow, thymus, heart, lymph, blood, bone, cartilage, pancreas, kidney, gall bladder, stomach, intestine, testis, ovary, uterus, rectum, nervous system, eye, gland, and connective tissue. Interstitial space of the tissues comprises the intercellular, fluid, mucopolysaccharide matrix among the reticular fibers of organ tissues, elastic fibers in the walls of vessels or chambers, collagen fibers of fibrous tissues, or that same matrix within connective tissue ensheathing muscle cells or in the lacunae of bone. It is similarly the space occupied by the plasma of the circulation and the lymph fluid of the lymphatic channels. Delivery to the interstitial space of muscle tissue is preferred for the reasons discussed below. They may be conveniently delivered by injection into the tissues comprising these cells. They are preferably delivered to and expressed in persistent, non-dividing cells which are differentiated, although delivery and expression may be achieved in non-differentiated or less completely differentiated cells, such as, for example, stem cells of blood or skin fibroblasts. *In vivo* muscle cells are particularly competent in their ability to take up and express polynucleotides.

For the naked acid sequence injection, an effective dosage amount of DNA or RNA will be in the range of from about 0.05 mg/kg body weight to about 50 mg/kg body weight. Preferably the dosage will be from about 0.005 mg/kg to about 20 mg/kg and more preferably from about 0.05 mg/kg to about 5 mg/kg. Of course, as the artisan of ordinary skill will appreciate, this dosage will vary according to the tissue site of injection. The appropriate and effective dosage of nucleic acid sequence can readily

be determined by those of ordinary skill in the art and may depend on the condition being treated and the route of administration.

The preferred route of administration is by the parenteral route of injection into the interstitial space or tissues. However, other parenteral routes may also be used, such as, inhalation of an aerosol formulation particularly for delivery to lungs or bronchial tissues, throat or mucous membranes of the nose. In addition, naked CRCGCL DNA constructs can be delivered to arteries during angioplasty by the catheter used in the procedure.

The naked polynucleotides are delivered by any method known in the art, including, but not limited to, direct needle injection at the delivery site, intravenous injection, topical administration, catheter infusion, and so-called "gene guns". These delivery methods are known in the art.

As is evidenced in the Examples, naked CRCGCL nucleic acid sequences can be administered *in vivo* results in the successful expression of CRCGCL polypeptide in the femoral arteries of rabbits.

The constructs may also be delivered with delivery vehicles such as viral sequences, viral particles, liposome formulations, lipofectin, precipitating agents, etc. Such methods of delivery are known in the art.

In certain embodiments, the CRCGCL polynucleotide constructs are complexed in a liposome preparation. Liposomal preparations for use in the instant invention include cationic (positively charged), anionic (negatively charged) and neutral preparations. However, cationic liposomes are particularly preferred because a tight charge complex can be formed between the cationic liposome and the polyanionic nucleic acid. Cationic liposomes have been shown to mediate intracellular delivery of plasmid DNA (Felgner *et al.*, *Proc. Natl. Acad. Sci. USA* (1987) 84:7413-7416, which is herein incorporated by reference); mRNA (Malone *et al.*, *Proc. Natl. Acad. Sci. USA* (1989) 86:6077-6081, which is herein incorporated by reference); and purified transcription factors (Debs *et al.*, *J. Biol. Chem.* (1990) 265:10189-10192, which is herein incorporated by reference), in functional form.

Cationic liposomes are readily available. For example, N[1-2,3-dioleoyloxy)propyl]-N,N,N-triethylammonium (DOTMA) liposomes are particularly useful and are available under the trademark Lipofectin, from GIBCO BRL, Grand Island, N.Y. (See, also, Felgner *et al.*, *Proc. Natl. Acad. Sci. USA* (1987) 84:7413-7416, which is herein incorporated by reference). Other commercially available liposomes include transfectace (DDAB/DOPE) and DOTAP/DOPE (Boehringer).

Other cationic liposomes can be prepared from readily available materials using techniques well known in the art. See, e.g., PCT Publication No. WO 90/11092 (which is herein incorporated by reference) for a description of the synthesis of DOTAP (1,2-bis(oleoyloxy)-3-(trimethylammonio)propane) liposomes. Preparation of DOTMA liposomes is explained in the literature, see, e.g., P. Felgner *et al.*, *Proc. Natl. Acad. Sci. USA* 84:7413-7417, which is herein incorporated by reference. Similar methods can be used to prepare liposomes from other cationic lipid materials.

Similarly, anionic and neutral liposomes are readily available, such as from Avanti Polar Lipids (Birmingham, Ala.), or can be easily prepared using readily available materials. Such materials include phosphatidyl choline, cholesterol, phosphatidyl ethanolamine, dioleoylphosphatidyl choline (DOPC), dioleoylphosphatidyl glycerol (DOPG), dioleoylphosphatidyl ethanolamine (DOPE), among others. These materials can also be mixed with the DOTMA and DOTAP starting materials in appropriate ratios. Methods for making liposomes using these materials are well known in the art.

For example, commercially dioleoylphosphatidyl choline (DOPC), dioleoylphosphatidyl glycerol (DOPG), and dioleoylphosphatidyl ethanolamine (DOPE) can be used in various combinations to make conventional liposomes, with or without the addition of cholesterol. Thus, for example, DOPG/DOPC vesicles can be prepared by drying 50 mg each of DOPG and DOPC under a stream of nitrogen gas into a sonication vial. The sample is placed under a vacuum pump overnight and is hydrated the following day with deionized water. The sample is then sonicated for 2 hours in a capped vial, using a Heat Systems model 350 sonicator equipped with an inverted cup (bath type) probe at the maximum setting while the bath is circulated at 15EC. Alternatively, negatively charged vesicles can be prepared without sonication to produce multilamellar vesicles or by extrusion through nucleopore membranes to produce unilamellar vesicles of discrete size. Other methods are known and available to those of skill in the art.

The liposomes can comprise multilamellar vesicles (MLVs), small unilamellar vesicles (SUVs), or large unilamellar vesicles (LUVs), with SUVs being preferred. The various liposome-nucleic acid complexes are prepared using methods well known in the art. See, e.g., Straubinger *et al.*, *Methods of Immunology* (1983), 101:512-527, which is herein incorporated by reference. For example, MLVs containing nucleic acid can be prepared by depositing a thin film of phospholipid on the walls of a glass tube and subsequently hydrating with a solution of the material to be encapsulated. SUVs are prepared by extended sonication of MLVs to produce a homogeneous population of unilamellar liposomes. The material to be entrapped is

added to a suspension of preformed MLVs and then sonicated. When using liposomes containing cationic lipids, the dried lipid film is resuspended in an appropriate solution such as sterile water or an isotonic buffer solution such as 10 mM Tris/NaCl, sonicated, and then the preformed liposomes are mixed directly with the DNA. The liposome and DNA form a very stable complex due to binding of the positively charged liposomes to the cationic DNA. SUVs find use with small nucleic acid fragments. LUVs are prepared by a number of methods, well known in the art. Commonly used methods include  $\text{Ca}^{2+}$ -EDTA chelation (Papahadjopoulos *et al.*, *Biochim. Biophys. Acta* (1975) 394:483; Wilson *et al.*, *Cell* (1979) 17:77), ether injection (Deamer, D. and Bangham, A., *Biochim. Biophys. Acta* (1976) 443:629; Ostro *et al.*, *Biochem. Biophys. Res. Commun.* (1977) 76:836; Fraley *et al.*, *Proc. Natl. Acad. Sci. USA* (1979) 76:3348); detergent dialysis (Enoch, H. and Strittmatter, P., *Proc. Natl. Acad. Sci. USA* (1979) 76:145); and reverse-phase evaporation (REV) (Fraley *et al.*, *J. Biol. Chem.* (1980) 255:10431; Szoka, F. and Papahadjopoulos, D., *Proc. Natl. Acad. Sci. USA* (1978) 75:145; Schaefer-Ridder *et al.*, *Science* (1982) 215:166), which are herein incorporated by reference.

Generally, the ratio of DNA to liposomes will be from about 10:1 to about 1:10. Preferably, the ration will be from about 5:1 to about 1:5. More preferably, the ration will be about 3:1 to about 1:3. Still more preferably, the ratio will be about 1:1.

U.S. Patent No. 5,676,954 (which is herein incorporated by reference) reports on the injection of genetic material, complexed with cationic liposomes carriers, into mice. U.S. Patent Nos. 4,897,355, 4,946,787, 5,049,386, 5,459,127, 5,589,466, 5,693,622, 5,580,859, 5,703,055, and international publication no. WO 94/9469 (which are herein incorporated by reference) provide cationic lipids for use in transfecting DNA into cells and mammals. U.S. Patent Nos. 5,589,466, 5,693,622, 5,580,859, 5,703,055, and international publication no. WO 94/9469 (which are herein incorporated by reference) provide methods for delivering DNA-cationic lipid complexes to mammals.

In certain embodiments, cells are be engineered, *ex vivo* or *in vivo*, using a retroviral particle containing RNA which comprises a sequence encoding CRGCL. Retroviruses from which the retroviral plasmid vectors may be derived include, but are not limited to, Moloney Murine Leukemia Virus, spleen necrosis virus, Rous sarcoma Virus, Harvey Sarcoma Virus, avian leukosis virus, gibbon ape leukemia virus, human immunodeficiency virus, Myeloproliferative Sarcoma Virus, and mammary tumor virus.

The retroviral plasmid vector is employed to transduce packaging cell lines to form producer cell lines. Examples of packaging cells which may be transfected



include, but are not limited to, the PE501, PA317, R-2, R-AM, PA12, T19-14X, VT-19-17-H2, RCRE, RCRIP, GP+E 86, GP+envAm12, and DAN cell lines as described in Miller, *Human Gene Therapy* 1:5-14 (1990), which is incorporated herein by reference in its entirety. The vector may transduce the packaging cells through any means known in the art. Such means include, but are not limited to, electroporation, the use of liposomes, and CaPO<sub>4</sub> precipitation. In one alternative, the retroviral plasmid vector may be encapsulated into a liposome, or coupled to a lipid, and then administered to a host.

The producer cell line generates infectious retroviral vector particles which include polynucleotide encoding CRGCL. Such retroviral vector particles then may be employed, to transduce eukaryotic cells, either *in vitro* or *in vivo*. The transduced eukaryotic cells will express CRGCL.

In certain other embodiments, cells are engineered, *ex vivo* or *in vivo*, with CRGCL polynucleotide contained in an adenovirus vector. Adenovirus can be manipulated such that it encodes and expresses CRGCL, and at the same time is inactivated in terms of its ability to replicate in a normal lytic viral life cycle. Adenovirus expression is achieved without integration of the viral DNA into the host cell chromosome, thereby alleviating concerns about insertional mutagenesis. Furthermore, adenoviruses have been used as live enteric vaccines for many years with an excellent safety profile (Schwartz, A. R. *et al.* (1974) *Am. Rev. Respir. Dis.* 109:233-238). Finally, adenovirus mediated gene transfer has been demonstrated in a number of instances including transfer of alpha-1-antitrypsin and CFTR to the lungs of cotton rats (Rosenfeld, M. A. *et al.* (1991) *Science* 252:431-434; Rosenfeld *et al.*, (1992) *Cell* 68:143-155). Furthermore, extensive studies to attempt to establish adenovirus as a causative agent in human cancer were uniformly negative (Green, M. *et al.* (1979) *Proc. Natl. Acad. Sci. USA* 76:6606).

Suitable adenoviral vectors useful in the present invention are described, for example, in Kozarsky and Wilson, *Curr. Opin. Genet. Devel.* 3:499-503 (1993); Rosenfeld *et al.*, *Cell* 68:143-155 (1992); Engelhardt *et al.*, *Human Genet. Ther.* 4:759-769 (1993); Yang *et al.*, *Nature Genet.* 7:362-369 (1994); Wilson *et al.*, *Nature* 365:691-692 (1993); and U.S. Patent No. 5,652,224, which are herein incorporated by reference. For example, the adenovirus vector Ad2 is useful and can be grown in human 293 cells. These cells contain the E1 region of adenovirus and constitutively express E1a and E1b, which complement the defective adenoviruses by providing the products of the genes deleted from the vector. In addition to Ad2, other varieties of adenovirus (e.g., Ad3, Ad5, and Ad7) are also useful in the present invention.

Preferably, the adenoviruses used in the present invention are replication deficient. Replication deficient adenoviruses require the aid of a helper virus and/or packaging cell line to form infectious particles. The resulting virus is capable of infecting cells and can express a polynucleotide of interest which is operably linked to a promoter, for example, the HARP promoter of the present invention, but cannot replicate in most cells. Replication deficient adenoviruses may be deleted in one or more of all or a portion of the following genes: E1a, E1b, E3, E4, E2a, or L1 through L5.

In certain other embodiments, the cells are engineered, *ex vivo* or *in vivo*, using an adeno-associated virus (AAV). AAVs are naturally occurring defective viruses that require helper viruses to produce infectious particles (Muzyczka, N., *Curr. Topics in Microbiol. Immunol.* 158:97 (1992)). It is also one of the few viruses that may integrate its DNA into non-dividing cells. Vectors containing as little as 300 base pairs of AAV can be packaged and can integrate, but space for exogenous DNA is limited to about 4.5 kb. Methods for producing and using such AAVs are known in the art. See, for example, U.S. Patent Nos. 5,139,941, 5,173,414, 5,354,678, 5,436,146, 5,474,935, 5,478,745, and 5,589,377.

For example, an appropriate AAV vector for use in the present invention will include all the sequences necessary for DNA replication, encapsidation, and host-cell integration. The CRGCL polynucleotide construct is inserted into the AAV vector using standard cloning methods, such as those found in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press (1989). The recombinant AAV vector is then transfected into packaging cells which are infected with a helper virus, using any standard technique, including lipofection, electroporation, calcium phosphate precipitation, etc. Appropriate helper viruses include adenoviruses, cytomegaloviruses, vaccinia viruses, or herpes viruses. Once the packaging cells are transfected and infected, they will produce infectious AAV viral particles which contain the CRGCL polynucleotide construct. These viral particles are then used to transduce eukaryotic cells, either *ex vivo* or *in vivo*. The transduced cells will contain the CRGCL polynucleotide construct integrated into its genome, and will express CRGCL.

Another method of gene therapy involves operably associating heterologous control regions and endogenous polynucleotide sequences (e.g. encoding CRGCL) via homologous recombination (see, e.g., U.S. Patent No. 5,641,670, issued June 24, 1997; International Publication No. WO 96/29411, published September 26, 1996; International Publication No. WO 94/12650, published August 4, 1994; Koller *et al.*, *Proc. Natl. Acad. Sci. USA* 86:8932-8935 (1989); and Zijlstra *et al.*, *Nature*

342,435-438 (1989). This method involves the activation of a gene which is present in the target cells, but which is not normally expressed in the cells, or is expressed at a lower level than desired.

5 Polynucleotide constructs are made, using standard techniques known in the art, which contain the promoter with targeting sequences flanking the promoter. Suitable promoters are described herein. The targeting sequence is sufficiently complementary to an endogenous sequence to permit homologous recombination of the promoter-targeting sequence with the endogenous sequence. The targeting  
10 sequence will be sufficiently near the 5' end of the CRGCL desired endogenous polynucleotide sequence so the promoter will be operably linked to the endogenous sequence upon homologous recombination.

The promoter and the targeting sequences can be amplified using PCR. Preferably, the amplified promoter contains distinct restriction enzyme sites on the 5' and 3' ends. Preferably, the 3' end of the first targeting sequence contains the same  
15 restriction enzyme site as the 5' end of the amplified promoter and the 5' end of the second targeting sequence contains the same restriction site as the 3' end of the amplified promoter. The amplified promoter and targeting sequences are digested and ligated together.

The promoter-targeting sequence construct is delivered to the cells, either as  
20 naked polynucleotide, or in conjunction with transfection-facilitating agents, such as liposomes, viral sequences, viral particles, whole viruses, lipofection, precipitating agents, etc., described in more detail above. The P promoter-targeting sequence can be delivered by any method, included direct needle injection, intravenous injection, topical administration, catheter infusion, particle accelerators, etc. The methods are  
25 described in more detail below.

The promoter-targeting sequence construct is taken up by cells. Homologous recombination between the construct and the endogenous sequence takes place, such that an endogenous CRGCL sequence is placed under the control of the promoter. The promoter then drives the expression of the endogenous CRGCL sequence.

30 The polynucleotides encoding CRGCL may be administered along with other polynucleotides encoding other angiogenic proteins. Angiogenic proteins include, but are not limited to, acidic and basic fibroblast growth factors, VEGF-1, epidermal growth factor alpha and beta, platelet-derived endothelial cell growth factor, platelet-derived growth factor, tumor necrosis factor alpha, hepatocyte growth factor, insulin  
35 like growth factor, colony stimulating factor, macrophage colony stimulating factor, granulocyte/macrophage colony stimulating factor, and nitric oxide synthase.

Preferably, the polynucleotide encoding CRGCL contains a secretory signal sequence that facilitates secretion of the protein. Typically, the signal sequence is positioned in the coding region of the polynucleotide to be expressed towards or at the 5' end of the coding region. The signal sequence may be homologous or heterologous to the polynucleotide of interest and may be homologous or heterologous to the cells to be transfected. Additionally, the signal sequence may be chemically synthesized using methods known in the art.

Any mode of administration of any of the above-described polynucleotides constructs can be used so long as the mode results in the expression of one or more molecules in an amount sufficient to provide a therapeutic effect. This includes direct needle injection, systemic injection, catheter infusion, biolistic injectors, particle accelerators (i.e., "gene guns"), gelfoam sponge depots, other commercially available depot materials, osmotic pumps (e.g., Alza minipumps), oral or suppositorial solid (tablet or pill) pharmaceutical formulations, and decanting or topical applications during surgery. For example, direct injection of naked calcium phosphate-precipitated plasmid into rat liver and rat spleen or a protein-coated plasmid into the portal vein has resulted in gene expression of the foreign gene in the rat livers (Kaneda *et al.*, *Science* 243:375 (1989)).

A preferred method of local administration is by direct injection. Preferably, a recombinant molecule of the present invention complexed with a delivery vehicle is administered by direct injection into or locally within the area of arteries. Administration of a composition locally within the area of arteries refers to injecting the composition centimeters and preferably, millimeters within arteries.

Another method of local administration is to contact a polynucleotide construct of the present invention in or around a surgical wound. For example, a patient can undergo surgery and the polynucleotide construct can be coated on the surface of tissue inside the wound or the construct can be injected into areas of tissue inside the wound.

Therapeutic compositions useful in systemic administration, include recombinant molecules of the present invention complexed to a targeted delivery vehicle of the present invention. Suitable delivery vehicles for use with systemic administration comprise liposomes comprising ligands for targeting the vehicle to a particular site.

Preferred methods of systemic administration, include intravenous injection, aerosol, oral and percutaneous (topical) delivery. Intravenous injections can be performed using methods standard in the art. Aerosol delivery can also be performed using methods standard in the art (see, for example, Stribling *et al.*, *Proc. Natl. Acad.*

*Sci. USA* 139:11277-11281, 1992, which is incorporated herein by reference). Oral delivery can be performed by complexing a polynucleotide construct of the present invention to a carrier capable of withstanding degradation by digestive enzymes in the gut of an animal. Examples of such carriers, include plastic capsules or tablets, such as those known in the art. Topical delivery can be performed by mixing a polynucleotide construct of the present invention with a lipophilic reagent (e.g., DMSO) that is capable of passing into the skin.

Determining an effective amount of substance to be delivered can depend upon a number of factors including, for example, the chemical structure and biological activity of the substance, the age and weight of the animal, the precise condition requiring treatment and its severity, and the route of administration. The frequency of treatments depends upon a number of factors, such as the amount of polynucleotide constructs administered per dose, as well as the health and history of the subject. The precise amount, number of doses, and timing of doses will be determined by the attending physician or veterinarian.

Therapeutic compositions of the present invention can be administered to any animal, preferably to mammals and birds. Preferred mammals include humans, dogs, cats, mice, rats, rabbits, sheep, cattle, horses and pigs, with humans being particularly preferred.

### **Biological Activities of CRCGCL**

CRCGCL polynucleotides and polypeptides can be used in assays to test for one or more biological activities. If CRCGCL polynucleotides and polypeptides do exhibit activity in a particular assay, it is likely that CRCGCL may be involved in the diseases associated with the biological activity. Therefore, CRCGCL could be used to treat the associated disease.

### **Immune Activity**

CRCGCL polypeptides or polynucleotides may be useful in treating deficiencies or disorders of the immune system, by activating or inhibiting the proliferation, differentiation, or mobilization (chemotaxis) of immune cells. Immune cells develop through a process called hematopoiesis, producing myeloid (platelets, red blood cells, neutrophils, and macrophages) and lymphoid (B and T lymphocytes) cells from pluripotent stem cells. The etiology of these immune deficiencies or disorders may be genetic, somatic, such as cancer or some autoimmune disorders, acquired (e.g., by chemotherapy or toxins), or infectious. Moreover, CRCGCL

polynucleotides or polypeptides can be used as a marker or detector of a particular immune system disease or disorder.

Interestingly, CRGCL maps to the pseudoautosomal regions on the X and Y chromosomes. It is likely that mutations in CRGCL may also lead to immune disorders, especially those involving activated T cells. Moreover, mutations in CRGCL may be involved in autoimmune diseases, especially X-linked autoimmune diseases.

CRGCL polynucleotides or polypeptides may be useful in treating or detecting deficiencies or disorders of hematopoietic cells. CRGCL polypeptides or polynucleotides could be used to increase differentiation and proliferation of hematopoietic cells, including the pluripotent stem cells, in an effort to treat those disorders associated with a decrease in certain (or many) types hematopoietic cells. Examples of immunologic deficiency syndromes include, but are not limited to: blood protein disorders (e.g., agammaglobulinemia, dysgammaglobulinemia), ataxia telangiectasia, common variable immunodeficiency, DiGeorge Syndrome, HIV infection, HTLV-BLV infection, leukocyte adhesion deficiency syndrome, lymphopenia, phagocyte bactericidal dysfunction, severe combined immunodeficiency (SCIDs), Wiskott-Aldrich Disorder, anemia, thrombocytopenia, or hemoglobinuria.

Moreover, CRGCL polypeptides or polynucleotides can also be used to modulate hemostatic (the stopping of bleeding) or thrombolytic activity (clot formation). For example, by increasing hemostatic or thrombolytic activity, CRGCL polynucleotides or polypeptides could be used to treat blood coagulation disorders (e.g., afibrinogenemia, factor deficiencies), blood platelet disorders (e.g., thrombocytopenia), or wounds resulting from trauma, surgery, or other causes. Alternatively, CRGCL polynucleotides or polypeptides that can decrease hemostatic or thrombolytic activity could be used to inhibit or dissolve clotting, important in the treatment of heart attacks (infarction), strokes, or scarring.

CRGCL polynucleotides or polypeptides may also be useful in treating or detecting autoimmune disorders. Many autoimmune disorders result from inappropriate recognition of self as foreign material by immune cells. This inappropriate recognition results in an immune response leading to the destruction of the host tissue. Therefore, the administration of CRGCL polypeptides or polynucleotides that can inhibit an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing autoimmune disorders.

Examples of autoimmune disorders that can be treated or detected by CRGCL include, but are not limited to Addison's Disease, hemolytic anemia,

antiphospholipid syndrome, rheumatoid arthritis, dermatitis, allergic encephalomyelitis, glomerulonephritis, Goodpasture's Syndrome, Graves' Disease, Multiple Sclerosis, Myasthenia Gravis, Neuritis, Ophthalmia, Bullous Pemphigoid, Penicilliosis, Polyendocrinopathies, Purpura, Reiter's Disease, Stiff-Man Syndrome, 5 Autoimmune Thyroiditis, Systemic Lupus Erythematosus, Autoimmune Pulmonary Inflammation, Guillain-Barre Syndrome, insulin dependent diabetes mellitus, and autoimmune inflammatory eye disease.

Similarly, allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems, may also be treated by CRCGCL 10 polypeptides or polynucleotides. Moreover, CRCGCL can be used to treat anaphylaxis, hypersensitivity to an antigenic molecule, or blood group incompatibility.

CRCGCL polynucleotides or polypeptides may also be used to treat and/or 15 prevent organ rejection or graft-versus-host disease (GVHD). Organ rejection occurs by host immune cell destruction of the transplanted tissue through an immune response. Similarly, an immune response is also involved in GVHD, but, in this case, the foreign transplanted immune cells destroy the host tissues. The administration of CRCGCL polypeptides or polynucleotides that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may 20 be an effective therapy in preventing organ rejection or GVHD.

Similarly, CRCGCL polypeptides or polynucleotides may also be used to modulate inflammation. For example, CRCGCL polypeptides or polynucleotides may inhibit the proliferation and differentiation of cells involved in an inflammatory response. These molecules can be used to treat inflammatory conditions, both chronic 25 and acute conditions, including inflammation associated with infection (e.g., septic shock, sepsis, or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine induced lung injury, inflammatory bowel disease, Crohn's disease, or resulting from over production of cytokines (e.g., TNF or IL-1.) 30

### **Hyperproliferative Disorders**

CRCGCL polypeptides or polynucleotides can be used to treat or detect hyperproliferative disorders, including neoplasms. CRCGCL polypeptides or polynucleotides may inhibit the proliferation of the disorder through direct or indirect interactions. Alternatively, CRCGCL polypeptides or polynucleotides may proliferate other cells which can inhibit the hyperproliferative disorder.

For example, by increasing an immune response, particularly increasing antigenic qualities of the hyperproliferative disorder or by proliferating, differentiating, or mobilizing T cells, hyperproliferative disorders can be treated. This immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, decreasing an immune response may also be a method of treating hyperproliferative disorders, such as a chemotherapeutic agent.

Examples of hyperproliferative disorders that can be treated or detected by CRCGCL polynucleotides or polypeptides include, but are not limited to neoplasms located in the: abdomen, bone, breast, digestive system, liver, pancreas, peritoneum, endocrine glands (adrenal, parathyroid, pituitary, testicles, ovary, thymus, thyroid), eye, head and neck, nervous (central and peripheral), lymphatic system, pelvic, skin, soft tissue, spleen, thoracic, and urogenital.

Similarly, other hyperproliferative disorders can also be treated or detected by CRCGCL polynucleotides or polypeptides. Examples of such hyperproliferative disorders include, but are not limited to: hypergammaglobulinemia, lymphoproliferative disorders, paraproteinemas, purpura, sarcoidosis, Sezary Syndrome, Waldenstrom's Macroglobulinemia, Gaucher's Disease, histiocytosis, and any other hyperproliferative disease, besides neoplasia, located in an organ system listed above.

### **Infectious Disease**

CRCGCL polypeptides or polynucleotides can be used to treat or detect infectious agents. For example, by increasing the immune response, particularly increasing the proliferation and differentiation of B and/or T cells, infectious diseases may be treated. The immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, CRCGCL polypeptides or polynucleotides may also directly inhibit the infectious agent, without necessarily eliciting an immune response.

Viruses are one example of an infectious agent that can cause disease or symptoms that can be treated or detected by CRCGCL polynucleotides or



polypeptides. Examples of viruses, include, but are not limited to the following DNA and RNA viral families: Arbovirus, Adenoviridae, Arenaviridae, Arterivirus, Birnaviridae, Bunyaviridae, Caliciviridae, Circoviridae, Coronaviridae, Flaviviridae, Hepadnaviridae (Hepatitis), Herpesviridae (such as, Cytomegalovirus, Herpes Simplex, Herpes Zoster), Mononegavirus (e.g., Paramyxoviridae, Morbillivirus, Rhabdoviridae), Orthomyxoviridae (e.g., Influenza), Papovaviridae, Parvoviridae, Picornaviridae, Poxviridae (such as, Smallpox or Vaccinia), Reoviridae (e.g., Rotavirus), Retroviridae (HTLV I, HTLV II, Lentivirus), and Togaviridae (e.g., Rubivirus). Viruses falling within these families can cause a variety of diseases or symptoms, including, but not limited to: arthritis, bronchiolitis, encephalitis, eye infections (e.g., conjunctivitis, keratitis), chronic fatigue syndrome, hepatitis (A, B, C, E, Chronic Active, Delta), meningitis, opportunistic infections (e.g., AIDS), pneumonia, Burkitt's Lymphoma, chickenpox, hemorrhagic fever, Measles, Mumps, Parainfluenza, Rabies, the common cold, Polio, leukemia, Rubella, sexually transmitted diseases, skin diseases (e.g., Kaposi's, warts), and viremia. CRCGCL polypeptides or polynucleotides can be used to treat or detect any of these symptoms or diseases.

Similarly, bacterial or fungal agents that can cause disease or symptoms and that can be treated or detected by CRCGCL polynucleotides or polypeptides include, but not limited to, the following Gram-Negative and Gram-positive bacterial families and fungi: Actinomycetales (e.g., Corynebacterium, Mycobacterium, Nocardia), Aspergillosis, Bacillaceae (e.g., Anthrax, Clostridium), Bacteroidaceae, Blastomycosis, Bordetella, Borrelia, Brucellosis, Candidiasis, Campylobacter, Coccidioidomycosis, Cryptococcosis, Dermatocycoses, Enterobacteriaceae (Klebsiella, Salmonella, Serratia, Yersinia), Erysipelothrix, Helicobacter, Legionellosis, Leptospirosis, Listeria, Mycoplasmatales, Neisseriaceae (e.g., Acinetobacter, Gonorrhea, Meningococcal), Pasteurellaceae Infections (e.g., Actinobacillus, Haemophilus, Pasteurella), Pseudomonas, Rickettsiaceae, Chlamydiaceae, Syphilis, and Staphylococcal. These bacterial or fungal families can cause the following diseases or symptoms, including, but not limited to: bacteremia, endocarditis, eye infections (conjunctivitis, tuberculosis, uveitis), gingivitis, opportunistic infections (e.g., AIDS related infections), paronychia, prosthesis-related infections, Reiter's Disease, respiratory tract infections, such as Whooping Cough or Empyema, sepsis, Lyme Disease, Cat-Scratch Disease, Dysentery, Paratyphoid Fever, food poisoning, Typhoid, pneumonia, Gonorrhea, meningitis, Chlamydia, Syphilis, Diphtheria, Leprosy, Paratuberculosis, Tuberculosis, Lupus, Botulism, gangrene, tetanus, impetigo, Rheumatic Fever, Scarlet Fever, sexually transmitted

diseases, skin diseases (e.g., cellulitis, dermatocycoses), toxemia, urinary tract infections, wound infections. CRCGCL polypeptides or polynucleotides can be used to treat or detect any of these symptoms or diseases.

Moreover, parasitic agents causing disease or symptoms that can be treated or  
5 detected by CRCGCL polynucleotides or polypeptides include, but not limited to, the following families: Amebiasis, Babesiosis, Coccidiosis, Cryptosporidiosis, Dientamoebiasis, Dourine, Ectoparasitic, Giardiasis, Helminthiasis, Leishmaniasis, Theileriasis, Toxoplasmosis, Trypanosomiasis, and Trichomonas. These parasites can cause a variety of diseases or symptoms, including, but not limited to: Scabies,  
10 Trombiculiasis, eye infections, intestinal disease (e.g., dysentery, giardiasis), liver disease, lung disease, opportunistic infections (e.g., AIDS related), Malaria, pregnancy complications, and toxoplasmosis. CRCGCL polypeptides or polynucleotides can be used to treat or detect any of these symptoms or diseases.

Preferably, treatment using CRCGCL polypeptides or polynucleotides could  
15 either be by administering an effective amount of CRCGCL polypeptide to the patient, or by removing cells from the patient, supplying the cells with CRCGCL polynucleotide, and returning the engineered cells to the patient (ex vivo therapy). Moreover, the CRCGCL polypeptide or polynucleotide can be used as an antigen in a vaccine to raise an immune response against infectious disease.

20

### **Regeneration**

CRCGCL polynucleotides or polypeptides can be used to differentiate, proliferate, and attract cells, leading to the regeneration of tissues. (See, Science 276:59-87 (1997).) The regeneration of tissues could be used to repair, replace, or  
25 protect tissue damaged by congenital defects, trauma (wounds, burns, incisions, or ulcers), age, disease (e.g. osteoporosis, osteoarthritis, periodontal disease, liver failure), surgery, including cosmetic plastic surgery, fibrosis, reperfusion injury, or systemic cytokine damage.

Tissues that could be regenerated using the present invention include organs  
30 (e.g., pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac), vasculature (including vascular and lymphatics), nervous, hematopoietic, and skeletal (bone, cartilage, tendon, and ligament) tissue. Preferably, regeneration occurs without or decreased scarring. Regeneration also may include angiogenesis.

Moreover, CRCGCL polynucleotides or polypeptides may increase  
35 regeneration of tissues difficult to heal. For example, increased tendon/ligament regeneration would quicken recovery time after damage. CRCGCL polynucleotides or polypeptides of the present invention could also be used prophylactically in an effort

to avoid damage. Specific diseases that could be treated include of tendinitis, carpal tunnel syndrome, and other tendon or ligament defects. A further example of tissue regeneration of non-healing wounds includes pressure ulcers, ulcers associated with vascular insufficiency, surgical, and traumatic wounds.

5 Similarly, nerve and brain tissue could also be regenerated by using CRCGCL polynucleotides or polypeptides to proliferate and differentiate nerve cells. Diseases that could be treated using this method include central and peripheral nervous system diseases, neuropathies, or mechanical and traumatic disorders (e.g., spinal cord disorders, head trauma, cerebrovascular disease, and stroke). Specifically, diseases  
10 associated with peripheral nerve injuries, peripheral neuropathy (e.g., resulting from chemotherapy or other medical therapies), localized neuropathies, and central nervous system diseases (e.g., Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome), could all be treated using the CRCGCL polynucleotides or polypeptides.

15

#### Chemotaxis

CRCGCL polynucleotides or polypeptides may have chemotaxis activity. A chemotaxic molecule attracts or mobilizes cells (e.g., monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells) to a  
20 particular site in the body, such as inflammation, infection, or site of hyperproliferation. The mobilized cells can then fight off and/or heal the particular trauma or abnormality.

CRCGCL polynucleotides or polypeptides may increase chemotaxic activity of particular cells. These chemotactic molecules can then be used to treat inflammation,  
25 infection, hyperproliferative disorders, or any immune system disorder by increasing the number of cells targeted to a particular location in the body. For example, chemotaxic molecules can be used to treat wounds and other trauma to tissues by attracting immune cells to the injured location. As a chemotactic molecule, CRCGCL could also attract fibroblasts, which can be used to treat wounds.

30 It is also contemplated that CRCGCL polynucleotides or polypeptides may inhibit chemotactic activity. These molecules could also be used to treat disorders. Thus, CRCGCL polynucleotides or polypeptides could be used as an inhibitor of chemotaxis.

35

#### Binding Activity

CRCGCL polypeptides may be used to screen for molecules that bind to CRCGCL or for molecules to which CRCGCL binds. The binding of CRCGCL and

the molecule may activate (agonist), increase, inhibit (antagonist), or decrease activity of the CRCGCL or the molecule bound. Examples of such molecules include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

5 Preferably, the molecule is closely related to the natural ligand of CRCGCL, e.g., a fragment of the ligand, or a natural substrate, a ligand, a structural or functional mimetic. (See, Coligan et al., Current Protocols in Immunology 1.2:Chapter 5 (1991).) Similarly, the molecule can be closely related to the natural receptor to which CRCGCL binds, or at least, a fragment of the receptor capable of being bound by CRCGCL (e.g., active site). In either case, the molecule can be  
10 rationally designed using known techniques.

Preferably, the screening for these molecules involves producing appropriate cells which express CRCGCL, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, *Drosophila*, or *E. coli*. Cells expressing CRCGCL (or cell membrane containing the expressed polypeptide) are then  
15 preferably contacted with a test compound potentially containing the molecule to observe binding, stimulation, or inhibition of activity of either CRCGCL or the molecule.

The assay may simply test binding of a candidate compound to CRCGCL, wherein binding is detected by a label, or in an assay involving competition with a  
20 labeled competitor. Further, the assay may test whether the candidate compound results in a signal generated by binding to CRCGCL.

Alternatively, the assay can be carried out using cell-free preparations, polypeptide/molecule affixed to a solid support, chemical libraries, or natural product mixtures. The assay may also simply comprise the steps of mixing a candidate  
25 compound with a solution containing CRCGCL, measuring CRCGCL/molecule activity or binding, and comparing the CRCGCL/molecule activity or binding to a standard.

Preferably, an ELISA assay can measure CRCGCL level or activity in a sample (e.g., biological sample) using a monoclonal or polyclonal antibody. The  
30 antibody can measure CRCGCL level or activity by either binding, directly or indirectly, to CRCGCL or by competing with CRCGCL for a substrate.

All of these above assays can be used as diagnostic or prognostic markers. The molecules discovered using these assays can be used to treat disease or to bring about a particular result in a patient (e.g., blood vessel growth) by activating or  
35 inhibiting the CRCGCL/molecule. Moreover, the assays can discover agents which may inhibit or enhance the production of CRCGCL from suitably manipulated cells or tissues.

Therefore, the invention includes a method of identifying compounds which bind to CRCGCL comprising the steps of: (a) incubating a candidate binding compound with CRCGCL; and (b) determining if binding has occurred. Moreover, the invention includes a method of identifying agonists/antagonists comprising the steps of: (a) incubating a candidate compound with CRCGCL; (b) assaying a biological activity; and (c) determining if a biological activity of CRCGCL has been altered.

#### Antisense And Ribozyme (Antagonists)

10 In specific embodiments, antagonists according to the present invention are nucleic acids corresponding to the sequences contained in SEQ ID NO.1, or the complementary strand thereof, and/or to nucleotide sequences contained in the deposited clone 209641 or 209691. In one embodiment, antisense sequence is generated internally by the organism. In another embodiment, the antisense sequence  
15 is separately administered (see, for example, O'Connor, J., Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988). Antisense technology can be used to control gene expression through antisense DNA or RNA, or through triple-helix formation. Antisense techniques are discussed for example, in Okano, J., Neurochem. 56:560  
20 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988). Triple helix formation is discussed in, for instance, Lee et al., Nucleic Acids Research 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1300 (1991). The methods are based on binding of a polynucleotide to a complementary DNA or RNA.

25 For example, the 5' coding portion of a polynucleotide that encodes the mature polypeptide of the present invention may be used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription thereby preventing transcription and the production of the receptor. The antisense RNA  
30 oligonucleotide hybridizes to the mRNA in vivo and blocks translation of the mRNA molecule into receptor polypeptide.

In one embodiment, the CRCGCL antisense nucleic acid of the invention is produced intracellularly by transcription from an exogenous sequence. For example, a vector or a portion thereof, is transcribed, producing an antisense nucleic acid  
35 (RNA) of the invention. Such a vector would contain a sequence encoding the CRCGCL antisense nucleic acid. Such a vector can remain episomal or become chromosomally integrated, as long as it can be transcribed to produce the desired

antisense RNA. Such vectors can be constructed by recombinant DNA technology methods standard in the art. Vectors can be plasmid, viral, or others known in the art, used for replication and expression in vertebrate cells. Expression of the sequence encoding CRCGCL, or fragments thereof, can be by any promoter known in the art to act in vertebrate, preferably human cells. Such promoters can be inducible or constitutive. Such promoters include, but are not limited to, the SV40 early promoter region (Bernorist and Chambon, *Nature* 29:304-310 (1981), the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto et al., *Cell* 22:787-797 (1980), the herpes thymidine promoter (Wagner et al., *Proc. Natl. Acad. Sci. U.S.A.* 78:1441-1445 (1981), the regulatory sequences of the metallothionein gene (Brinster, et al., *Nature* 296:39-42 (1982)), etc.

The antisense nucleic acids of the invention comprise a sequence complementary to at least a portion of an RNA transcript of a CRCGCL gene. However, absolute complementarity, although preferred, is not required. A sequence "complementary to at least a portion of an RNA," referred to herein, means a sequence having sufficient complementarity to be able to hybridize with the RNA, forming a stable duplex; in the case of double stranded CRCGCL antisense nucleic acids, a single strand of the duplex DNA may thus be tested, or triplex formation may be assayed. The ability to hybridize will depend on both the degree of complementarity and the length of the antisense nucleic acid. Generally, the larger the hybridizing nucleic acid, the more base mismatches with a CRCGCL RNA it may contain and still form a stable duplex (or triplex as the case may be). One skilled in the art can ascertain a tolerable degree of mismatch by use of standard procedures to determine the melting point of the hybridized complex.

Oligonucleotides that are complementary to the 5' end of the message, *e.g.*, the 5' untranslated sequence up to and including the AUG initiation codon, should work most efficiently at inhibiting translation. However, sequences complementary to the 3' untranslated sequences of mRNAs have been shown to be effective at inhibiting translation of mRNAs as well. See generally, Wagner, R., 1994, *Nature* 372:333-335. Thus, oligonucleotides complementary to either the 5'- or 3'- non- translated, non-coding regions of CRCGCL shown in Figures 1A-1B could be used in an antisense approach to inhibit translation of endogenous CRCGCL mRNA. Oligonucleotides complementary to the 5' untranslated region of the mRNA should include the complement of the AUG start codon. Antisense oligonucleotides complementary to mRNA coding regions are less efficient inhibitors of translation but could be used in accordance with the invention. Whether designed to hybridize to the 5'-, 3'- or coding region of CRCGCL mRNA, antisense nucleic acids should be at

least six nucleotides in length, and are preferably oligonucleotides ranging from 6 to about 50 nucleotides in length. In specific aspects the oligonucleotide is at least 10 nucleotides, at least 17 nucleotides, at least 25 nucleotides or at least 50 nucleotides.

The polynucleotides of the invention can be DNA or RNA or chimeric mixtures or derivatives or modified versions thereof, single-stranded or double-stranded. The oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone, for example, to improve stability of the molecule, hybridization, etc. The oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, e.g., Letsinger et al., 1989, Proc. Natl. Acad. Sci. U.S.A. 86:6553-6556; Lemaitre et al., 1987, Proc. Natl. Acad. Sci. 84:648-652, PCT Publication No. WO88/09810, published December 15, 1988) or the blood brain barrier (see, e.g., PCT Publication No. WO89/10134, published April 25, 1988), hybridization-triggered cleavage agents, (See, e.g., Krol et al., 1988, BioTechniques 6:958-976) or intercalating agents, (See, e.g., Zon, 1988, Pharm. Res. 5:539-549). To this end, the oligonucleotide may be conjugated to another molecule, e.g., a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

The antisense oligonucleotide may comprise at least one modified base moiety which is selected from the group including, but not limited to, 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

The antisense oligonucleotide may also comprise at least one modified sugar moiety selected from the group including, but not limited to, arabinose, 2-fluoroarabinose, xylulose, and hexose.

In yet another embodiment, the antisense oligonucleotide comprises at least one modified phosphate backbone selected from the group including, but not limited to, a phosphorothioate, a phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal or analog thereof.

In yet another embodiment, the antisense oligonucleotide is an anomeric oligonucleotide. An anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual  $\beta$ -units, the strands run parallel to each other (Gautier et al., 1987, Nucl. Acids Res. 15:6625-6641). The oligonucleotide is a 2'-O-methylribonucleotide (Inoue et al., 1987, Nucl. Acids Res. 15:6131-6148), or a chimeric RNA-DNA analogue (Inoue et al., 1987, FEBS Lett. 215:327-330).

Polynucleotides of the invention may be synthesized by standard methods known in the art, e.g., by use of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). As examples, phosphorothioate oligonucleotides may be synthesized by the method of Stein et al. (1988, Nucl. Acids Res. 16:3209), methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (Sarin et al., 1988, Proc. Natl. Acad. Sci. U.S.A. 85:7448-7451), etc.

While antisense nucleotides complementary to the CRCGCL coding region sequence could be used, those complementary to the transcribed untranslated region are most preferred.

Potential antagonists according to the invention also include catalytic RNA, or a ribozyme (See, e.g., PCT International Publication WO 90/11364, published October 4, 1990; Sarver et al. Science 247:1222-1225 (1990)). While ribozymes that cleave mRNA at site specific recognition sequences can be used to destroy CRCGCL mRNAs, the use of hammerhead ribozymes is preferred. Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. The sole requirement is that the target mRNA have the following sequence of two bases: 5'-UG-3'. The construction and production of hammerhead ribozymes is well known in the art and is described more fully in Haseloff and Gerlach, Nature 334:585-591 (1988). There are numerous potential hammerhead ribozyme cleavage sites within the nucleotide sequence of CRCGCL (Figures 1A-1B). Preferably, the ribozyme is engineered so that the cleavage recognition site is located near the 5' end of the CRCGCL mRNA; i.e., to increase efficiency and minimize the intracellular accumulation of non-functional mRNA transcripts.



As in the antisense approach, the ribozymes of the invention can be composed of modified oligonucleotides (e.g., for improved stability, targeting, etc.) and should be delivered to cells which express CRGCL *in vivo*. DNA constructs encoding the ribozyme may be introduced into the cell in the same manner as described above for the introduction of antisense encoding DNA. A preferred method of delivery involves using a DNA construct "encoding" the ribozyme under the control of a strong constitutive promoter, such as, for example, pol III or pol II promoter, so that transfected cells will produce sufficient quantities of the ribozyme to destroy endogenous CRGCL messages and inhibit translation. Since ribozymes unlike antisense molecules, are catalytic, a lower intracellular concentration is required for efficiency.

#### Other Activities

CRGCL polypeptides or polynucleotides may also increase or decrease the differentiation or proliferation of embryonic stem cells, besides, as discussed above, hematopoietic lineage.

CRGCL polypeptides or polynucleotides may also be used to modulate mammalian characteristics, such as body height, weight, hair color, eye color, skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery). Similarly, CRGCL polypeptides or polynucleotides may be used to modulate mammalian metabolism affecting catabolism, anabolism, processing, utilization, and storage of energy.

CRGCL polypeptides or polynucleotides may be used to change a mammal's mental state or physical state by influencing biorhythms, cardiac rhythms, depression (including depressive disorders), tendency for violence, tolerance for pain, reproductive capabilities (preferably by Activin or Inhibin-like activity), hormonal or endocrine levels, appetite, libido, memory, stress, or other cognitive qualities.

CRGCL polypeptides or polynucleotides may also be used as a food additive or preservative, such as to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

## Examples

### Example 1: Isolation of the CRCGCL cDNA Clone From the Deposited Sample

5           The cDNA for CRCGCL is inserted into the EcoRI/XhoI multiple cloning site of Uni-ZAP XR (Stratagene). Uni-ZAP XR contains an ampicillin resistance gene and may be transformed into *E. coli* strain DH10B, available from Life Technologies (See, for instance, Gruber, C. E., et al., *Focus* 15: 59 (1993)).

10           Two approaches can be used to isolate CRCGCL from the deposited sample. First, the deposited clone is transformed into a suitable host (such as XL-1 Blue (Stratagene)) using techniques known to those of skill in the art, such as those provided by the vector supplier or in related publications or patents. The transformants are plated on 1.5% agar plates (containing the appropriate selection agent, e.g., ampicillin) to a density of about 150 transformants (colonies) per plate. A single colony is then used to generate DNA using nucleic acid isolation techniques well known to those skilled in the art (e.g., Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd Edit., (1989), Cold Spring Harbor Laboratory Press.)

20           Alternatively, two primers of 17-20 nucleotides derived from both ends of the SEQ ID NO:1 (i.e., within the region of SEQ ID NO:1 bounded by the 5' NT and the 3' NT of the clone) are synthesized and used to amplify the CRCGCL cDNA using the deposited cDNA plasmid as a template. The polymerase chain reaction is carried out under routine conditions, for instance, in 25 ul of reaction mixture with 0.5 ug of the above cDNA template. A convenient reaction mixture is 1.5-5 mM MgCl<sub>2</sub>, 0.01% (w/v) gelatin, 20 uM each of dATP, dCTP, dGTP, dTTP, 25 pmol of each primer and 0.25 Unit of Taq polymerase. Thirty five cycles of PCR (denaturation at 94 degree C for 1 min; annealing at 55 degree C for 1 min; elongation at 72 degree C for 1 min) are performed with a Perkin-Elmer Cetus automated thermal cycler. The amplified product is analyzed by agarose gel electrophoresis and the DNA band with expected molecular weight is excised and purified. The PCR product is verified to be the selected sequence by subcloning and sequencing the DNA product.

30           Several methods are available for the identification of the 5' or 3' non-coding portions of the CRCGCL gene which may not be present in the deposited clone. These methods include but are not limited to, filter probing, clone enrichment using specific probes, and protocols similar or identical to 5' and 3' "RACE" protocols which are well known in the art. For instance, a method similar to 5' RACE is

available for generating the missing 5' end of a desired full-length transcript. (Fromont-Racine et al., Nucleic Acids Res. 21(7):1683-1684 (1993).)

5 Briefly, a specific RNA oligonucleotide is ligated to the 5' ends of a population of RNA presumably containing full-length gene RNA transcripts. A primer set containing a primer specific to the ligated RNA oligonucleotide and a primer specific to a known sequence of the CRCGCL gene of interest is used to PCR amplify the 5' portion of the CRCGCL full-length gene. This amplified product may then be sequenced and used to generate the full length gene.

10 This above method starts with total RNA isolated from the desired source, although poly-A+ RNA can be used. The RNA preparation can then be treated with phosphatase if necessary to eliminate 5' phosphate groups on degraded or damaged RNA which may interfere with the later RNA ligase step. The phosphatase should then be inactivated and the RNA treated with tobacco acid pyrophosphatase in order to remove the cap structure present at the 5' ends of messenger RNAs. This reaction  
15 leaves a 5' phosphate group at the 5' end of the cap cleaved RNA which can then be ligated to an RNA oligonucleotide using T4 RNA ligase.

This modified RNA preparation is used as a template for first strand cDNA synthesis using a gene specific oligonucleotide. The first strand synthesis reaction is used as a template for PCR amplification of the desired 5' end using a primer specific  
20 to the ligated RNA oligonucleotide and a primer specific to the known sequence of the gene of interest. The resultant product is then sequenced and analyzed to confirm that the 5' end sequence belongs to the CRCGCL gene.

#### **Example 2: Isolation of CRCGCL Genomic Clones**

25 A human genomic P1 library (Genomic Systems, Inc.) is screened by PCR using primers selected for the cDNA sequence corresponding to SEQ ID NO:1., according to the method described in Example 1. (See also, Sambrook.)

#### **Example 3: Tissue Distribution of CRCGCL Polypeptides**

30 Tissue distribution of mRNA expression of CRCGCL is determined using protocols for Northern blot analysis, described by, among others, Sambrook et al. For example, a CRCGCL probe produced by the method described in Example 1 is labeled with P<sup>32</sup> using the rediprime™ DNA labeling system (Amersham Life Science), according to manufacturer's instructions. After labeling, the probe is  
35 purified using CHROMA SPIN-100™ column (Clontech Laboratories, Inc.), according to manufacturer's protocol number PT1200-1. The purified labeled probe is then used to examine various human tissues for mRNA expression.

Multiple Tissue Northern (MTN) blots containing various human tissues (H) or human immune system tissues (IM) (Clontech) are examined with the labeled probe using ExpressHyb<sup>TM</sup> hybridization solution (Clontech) according to manufacturer's protocol number PT1190-1. Following hybridization and washing, the blots are mounted and exposed to film at 70 degree C overnight, and the films developed according to standard procedures.

#### **Example 4: Chromosomal Mapping of CRCGCL**

An oligonucleotide primer set is designed according to the sequence at the 5' end of SEQ ID NO:1. This primer preferably spans about 100 nucleotides. This primer set is then used in a polymerase chain reaction under the following set of conditions: 30 seconds, 95 degree C; 1 minute, 56 degree C; 1 minute, 70 degree C. This cycle is repeated 32 times followed by one 5 minute cycle at 70 degree C. Human, mouse, and hamster DNA is used as template in addition to a somatic cell hybrid panel containing individual chromosomes or chromosome fragments (Bios, Inc.). The reactions is analyzed on either 8% polyacrylamide gels or 3.5 % agarose gels. Chromosome mapping is determined by the presence of an approximately 100 bp PCR fragment in the particular somatic cell hybrid.

#### **Example 5: Bacterial Expression of CRCGCL**

CRCGCL polynucleotide encoding a CRCGCL polypeptide invention is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' ends of the DNA sequence, as outlined in Example 1, to synthesize insertion fragments. The primers used to amplify the cDNA insert should preferably contain restriction sites, such as BamHI and XbaI, at the 5' end of the primers in order to clone the amplified product into the expression vector. For example, BamHI and XbaI correspond to the restriction enzyme sites on the bacterial expression vector pQE-9. (Qiagen, Inc., Chatsworth, CA). This plasmid vector encodes antibiotic resistance (Amp<sup>r</sup>), a bacterial origin of replication (ori), an IPTG-regulatable promoter/operator (P/O), a ribosome binding site (RBS), a 6-histidine tag (6-His), and restriction enzyme cloning sites.

Specifically, to clone the CRCGCL protein in a bacterial vector, the 5' primer has the sequence 5' gttagccatgggaggagcagcagaagga 3' (SEQ ID NO: 14) containing the Nco I restriction site followed a number of nucleotides of the amino terminal coding sequence of the CRCGCL sequence in SEQ ID NO:1. One of ordinary skill in the art would appreciate, of course, that the point in the protein coding sequence

where the 5' primer begins may be varied to amplify a DNA segment encoding any desired portion of the complete CRCGCL protein shorter or longer than the portion described above. The 3' primer has the sequence 5'-ggtaaagatctcaacgccacgtaggagcggtc-3' (SEQ ID NO:15) containing the BglII restriction site followed by a number nucleotides complementary to the 3' end of the coding sequence of the CRCGCL DNA sequence of SEQ ID NO:1.

The pQE-9 vector is digested with BamHI and XbaI and the amplified fragment is ligated into the pQE-9 vector maintaining the reading frame initiated at the bacterial RBS. The ligation mixture is then used to transform the E. coli strain M15/rep4 (Qiagen, Inc.) which contains multiple copies of the plasmid pREP4, which expresses the lacI repressor and also confers kanamycin resistance (Kan<sup>r</sup>). Transformants are identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies are selected. Plasmid DNA is isolated and confirmed by restriction analysis.

Clones containing the desired constructs are grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 ug/ml). The O/N culture is used to inoculate a large culture at a ratio of 1:100 to 1:250. The cells are grown to an optical density 600 (O.D.<sub>600</sub>) of between 0.4 and 0.6. IPTG (Isopropyl-B-D-thiogalacto pyranoside) is then added to a final concentration of 1 mM. IPTG induces by inactivating the lacI repressor, clearing the P/O leading to increased gene expression.

Cells are grown for an extra 3 to 4 hours. Cells are then harvested by centrifugation (20 mins at 6000Xg). The cell pellet is solubilized in the chaotropic agent 6 Molar Guanidine HCl by stirring for 3-4 hours at 4°C. The cell debris is removed by centrifugation, and the supernatant containing the polypeptide is loaded onto a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (available from QIAGEN, Inc., *supra*). Proteins with a 6 x His tag bind to the Ni-NTA resin with high affinity and can be purified in a simple one-step procedure (for details see: The QIAexpressionist (1995) QIAGEN, Inc., *supra*).

Briefly, the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8, the column is first washed with 10 volumes of 6 M guanidine-HCl, pH 8, then washed with 10 volumes of 6 M guanidine-HCl pH 6, and finally the polypeptide is eluted with 6 M guanidine-HCl, pH 5.

The purified CRCGCL protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the CRCGCL protein can be successfully refolded while

immobilized on the Ni-NTA column. The recommended conditions are as follows: renature using a linear 6M-1M urea gradient in 500 mM NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4, containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After renaturation the proteins are  
5 eluted by the addition of 250 mM imidazole. Imidazole is removed by a final dialyzing step against PBS or 50 mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified CRCGCL protein is stored at 4°C or frozen at -80°C.

In addition to the above expression vector, the present invention further includes an expression vector comprising phage operator and promoter elements  
10 operatively linked to a CRCGCL polynucleotide, called pHE4a. (ATCC Accession Number 209645, deposited February 25, 1998.) This vector contains: 1) a neomycinphosphotransferase gene as a selection marker, 2) an *E. coli* origin of replication, 3) a T5 phage promoter sequence, 4) two lac operator sequences, 5) a Shine-Delgarno sequence, and 6) the lactose operon repressor gene (*lacIq*). The  
15 origin of replication (*oriC*) is derived from pUC19 (LTI, Gaithersburg, MD). The promoter sequence and operator sequences are made synthetically.

DNA can be inserted into the pHEa by restricting the vector with NdeI and XbaI, BamHI, XhoI, or Asp718, running the restricted product on a gel, and isolating the larger fragment (the stuffer fragment should be about 310 base pairs). The DNA  
20 insert is generated according to the PCR protocol described in Example 1, using PCR primers having restriction sites for NdeI (5' primer) and XbaI, BamHI, XhoI, or Asp718 (3' primer). The PCR insert is gel purified and restricted with compatible enzymes. The insert and vector are ligated according to standard protocols.

The engineered vector could easily be substituted in the above protocol to  
25 express protein in a bacterial system. More preferably, the bacterial expression vector, pQE60 can also be used to express CRCGCL.

#### **Example 6: Purification of CRCGCL Polypeptide from an Inclusion** 30 **Body**

The following alternative method can be used to purify CRCGCL polypeptide expressed in *E. coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10 degree C.

Upon completion of the production phase of the *E. coli* fermentation, the cell  
35 culture is cooled to 4-10 degree C and the cells harvested by continuous centrifugation at 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an appropriate

amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

5 The cells are then lysed by passing the solution through a microfluidizer (Microfluidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 xg for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

10 The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 xg centrifugation for 15 min., the pellet is discarded and the polypeptide containing supernatant is incubated at 4 degree C overnight to allow further GuHCl extraction.

Following high speed centrifugation (30,000 xg) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl  
15 extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring. The refolded diluted protein solution is kept at 4 degree C without mixing for 12 hours prior to further purification steps.

To clarify the refolded polypeptide solution, a previously prepared tangential  
20 filtration unit equipped with 0.16  $\mu$ m membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a  
25 stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

Fractions containing the CRCGCL polypeptide are then pooled and mixed with 4 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins.  
30 The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant  $A_{280}$  monitoring of the effluent. Fractions  
35 containing the polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

The resultant CRCGCL polypeptide should exhibit greater than 95% purity after the above refolding and purification steps. No major contaminant bands should be observed from Commassie blue stained 16% SDS-PAGE gel when 5 ug of purified protein is loaded. The purified CRCGCL protein can also be tested for  
5 endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

#### **Example 7: Cloning and Expression of CRCGCL in a Baculovirus Expression System**

10 In this example, the plasmid shuttle vector pA2 is used to insert CRCGCL polynucleotide into a baculovirus to express CRCGCL. This expression vector contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites such as BamHI, Xba I and Asp718. The polyadenylation site of the simian virus 40 ("SV40")  
15 is used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from *E. coli* under control of a weak Drosophila promoter in the same orientation, followed by the polyadenylation signal of the polyhedrin gene. The inserted genes are flanked on both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to  
20 generate a viable virus that express the cloned CRCGCL polynucleotide.

Many other baculovirus vectors can be used in place of the vector above, such as pAc373, pVL941, and pAcIM1, as one skilled in the art would readily appreciate, as long as the construct provides appropriately located signals for transcription, translation, secretion and the like, including a signal peptide and an in-frame AUG as  
25 required. Such vectors are described, for instance, in Luckow et al., *Virology* 170:31-39 (1989).

Specifically, the CRCGCL cDNA sequence contained in the deposited clone, including the AUG initiation codon and any naturally associated leader sequence, is amplified using the PCR protocol described in Example 1. If the naturally occurring  
30 signal sequence is used to produce the secreted protein, the pA2 vector does not need a second signal peptide. Alternatively, the vector can be modified (pA2 GP) to include a baculovirus leader sequence, using the standard methods described in Summers et al., "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures," Texas Agricultural Experimental Station Bulletin No. 1555  
35 (1987).

More specifically, the cDNA sequence encoding the full length CRCGCL protein in the deposited clone, including the AUG initiation codon and the naturally



associated leader sequence shown in SEQ ID NO:1, is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' sequences of the gene. The 5' primer has the sequence 5' ceggttagatctgccatcatgctttggggcaaggagg 3' (SEQ ID NO: 16) containing the BglII restriction enzyme site, an efficient signal for initiation of translation in eukaryotic cells (Kozak, M., *J. Mol. Biol.* 196:947-950 (1987)), followed by a number of nucleotides of the sequence of the complete CRGCI protein shown in Figures 1A-1B, beginning with the AUG initiation codon. The 3' primer has the sequence 5' ceggtttctagatcacaaggecaagtaggagcggtc 3' (SEQ ID NO: 17) containing the XbaI restriction site followed by a number of nucleotides complementary to the 3' noncoding sequence in Figures 1A-1B.

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("GeneClean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The plasmid is digested with the corresponding restriction enzymes and optionally, can be dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit ("GeneClean" BIO 101 Inc., La Jolla, Ca.).

The fragment and the dephosphorylated plasmid are ligated together with T4 DNA ligase. *E. coli* HB101 or other suitable *E. coli* hosts such as XL-1 Blue (Stratagene Cloning Systems, La Jolla, CA) cells are transformed with the ligation mixture and spread on culture plates. Bacteria containing the plasmid are identified by digesting DNA from individual colonies and analyzing the digestion product by gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing.

Five ug of a plasmid containing the polynucleotide is co-transfected with 1.0 ug of a commercially available linearized baculovirus DNA ("BaculoGold™ baculovirus DNA", Pharmingen, San Diego, CA), using the lipofection method described by Felgner et al., *Proc. Natl. Acad. Sci. USA* 84:7413-7417 (1987). One ug of BaculoGold™ virus DNA and 5 ug of the plasmid are mixed in a sterile well of a microtiter plate containing 50 ul of serum-free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards, 10 ul Lipofectin plus 90 ul Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is then incubated for 5 hours at 27 degrees C. The transfection solution is then removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27 degrees C for four days.

After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, *supra*. An agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a plaque assay" of this type can also be found in the user's guide for insect cell culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10.) After appropriate incubation, blue stained plaques are picked with the tip of a micropipettor (e.g., Eppendorf). The agar containing the recombinant viruses is then resuspended in a microcentrifuge tube containing 200  $\mu$ l of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4 degree C.

To verify the expression of the polypeptide, Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus containing the polynucleotide at a multiplicity of infection ("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies Inc., Rockville, MD). After 42 hours, 5  $\mu$ Ci of  $^{35}$ S-methionine and 5  $\mu$ Ci  $^{35}$ S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then are harvested by centrifugation. The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the produced CRGCL protein.

#### **Example 8: Expression of CRGCL in Mammalian Cells**

CRGCL polypeptide can be expressed in a mammalian cell. A typical mammalian expression vector contains a promoter element, which mediates the initiation of transcription of mRNA, a protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription is achieved with the early and late promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLVI, HIVI and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter).

Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2DHFR (ATCC 37146), pBC12MI (ATCC 67109), pCMVSPORT 2.0, and pCMVSPORT 3.0. Mammalian host cells that could be used  
5 include, human HeLa, 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

Alternatively, CRCGCL polypeptide can be expressed in stable cell lines containing the CRCGCL polynucleotide integrated into a chromosome. The co-  
10 transfection with a selectable marker such as DHFR, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

The transfected CRCGCL gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihydrofolate reductase) marker is useful in developing cell lines that carry several hundred or even several thousand copies of the gene of interest. (See, e.g., Alt, F. W., et al., *J. Biol. Chem.* 253:1357-1370 (1978);  
15 Handlin, J. L. and Ma, C., *Biochem. et Biophys. Acta*, 1097:107-143 (1990); Page, M. J. and Sydenham, M. A., *Biotechnology* 9:64-68 (1991).) Another useful selection marker is the enzyme glutamine synthase (GS) (Murphy et al., *Biochem J.* 227:277-279 (1991); Bebbington et al., *Bio/Technology* 10:169-175 (1992). Using  
20 these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the production of proteins.

Derivatives of the plasmid pSV2-DHFR (ATCC Accession No. 37146), the  
25 expression vectors pC4 (ATCC Accession No. 209646) and pC6 (ATCC Accession No. 209647) contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen et al., *Molecular and Cellular Biology*, 438-447 (March, 1985)) plus a fragment of the CMV-enhancer (Boshart et al., *Cell* 41:521-530 (1985)). Multiple cloning sites, e.g.,  
30 with the restriction enzyme cleavage sites BamHI, XbaI and Asp718, facilitate the cloning of CRCGCL. The vectors also contain the 3' intron, the polyadenylation and termination signal of the rat preproinsulin gene, and the mouse DHFR gene under control of the SV40 early promoter.

Specifically, the plasmid pC4, for example, is digested with appropriate restriction enzymes and then dephosphorylated using calf intestinal phosphates by  
35 procedures known in the art. The vector is then isolated from a 1% agarose gel. Also preferred is the pcDNA3 vector (Life Technologies).

CRCGCL polynucleotide is amplified according to the protocol outlined in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the vector does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

The amplified fragment is then digested with the same restriction enzyme and purified on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC4 using, for instance, restriction enzyme analysis.

Chinese hamster ovary cells lacking an active DHFR gene is used for transfection. Five µg of the expression plasmid pC4 is cotransfected with 0.5 µg of the plasmid pSVneo using lipofectin (Felgner et al., *supra*). The plasmid pSV2-neo contains a dominant selectable marker, the *neo* gene from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of methotrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1 µM, 2 µM, 5 µM, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100 - 200 µM. Expression of CRCGCL is analyzed, for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.

#### **Example 9: Construction of N-Terminal and/or C-Terminal Deletion Mutants**

The following general approach may be used to clone a N-terminal or C-terminal deletion CRCGCL deletion mutant. Generally, two oligonucleotide primers of about 15-25 nucleotides are derived from the desired 5' and 3' positions of a polynucleotide of SEQ ID NO:1. The 5' and 3' positions of the primers are determined based on the desired CRCGCL polynucleotide fragment. An initiation and stop codon are added to the 5' and 3' primers respectively, if necessary, to express the CRCGCL polypeptide fragment encoded by the polynucleotide fragment. Preferred CRCGCL polynucleotide fragments are those encoding the N-terminal and C-terminal

deletion mutants disclosed above in the "Polynucleotide and Polypeptide Fragments" section of the Specification.

Additional nucleotides containing restriction sites to facilitate cloning of the CRCGCL polynucleotide fragment in a desired vector may also be added to the 5' and 3' primer sequences. The CRCGCL polynucleotide fragment is amplified from genomic DNA or from the deposited cDNA clone using the appropriate PCR oligonucleotide primers and conditions discussed herein or known in the art. The CRCGCL polypeptide fragments encoded by the CRCGCL polynucleotide fragments of the present invention may be expressed and purified in the same general manner as the full length polypeptides, although routine modifications may be necessary due to the differences in chemical and physical properties between a particular fragment and full length polypeptide.

As a means of exemplifying but not limiting the present invention the polynucleotide encoding the CRCGCL polypeptide fragment I-35 to F-276 is amplified and cloned as follows: A 5' primer is generated comprising a restriction enzyme site followed by an initiation codon in frame with the polynucleotide sequence encoding the N-terminal portion of the polypeptide fragment beginning with I-35. A complementary 3' primer is generated comprising a restriction enzyme site followed by a stop codon in frame with the polynucleotide sequence encoding C-terminal portion of the CRCGCL polypeptide fragment ending with F-276.

The amplified polynucleotide fragment and the expression vector are digested with restriction enzymes that recognize the sites in the primers. The digested polynucleotides are then ligated together. The CRCGCL polynucleotide fragment is inserted into the restricted expression vector, preferably in a manner that places the CRCGCL polypeptide fragment coding region downstream from the promoter. The ligation mixture is transformed into competent *E. coli* cells using standard procedures and as described in the Examples herein. Plasmid DNA is isolated from resistant colonies and the identity of the cloned DNA confirmed by restriction analysis, PCR and DNA sequencing.

#### **Example 10: Protein Fusions of CRCGCL**

CRCGCL polypeptides are preferably fused to other proteins. These fusion proteins can be used for a variety of applications. For example, fusion of CRCGCL polypeptides to His-tag, HA-tag, protein A, IgG domains, and maltose binding protein facilitates purification. (See Example 5; see also EP A 394,827; Trautnecker, et al., *Nature* 331:84-86 (1988).) Similarly, fusion to IgG-1, IgG-3, and albumin increases the half-life time in vivo. Nuclear localization signals fused to CRCGCL

polypeptides can target the protein to a specific subcellular localization, while covalent heterodimer or homodimers can increase or decrease the activity of a fusion protein. Fusion proteins can also create chimeric molecules having more than one function. Finally, fusion proteins can increase solubility and/or stability of the fused protein compared to the non-fused protein. All of the types of fusion proteins described above can be made by modifying the following protocol, which outlines the fusion of a polypeptide to an IgG molecule, or the protocol described in Example 5.

Briefly, the human Fc portion of the IgG molecule can be PCR amplified, using primers that span the 5' and 3' ends of the sequence described below. These primers also should have convenient restriction enzyme sites that will facilitate cloning into an expression vector, preferably a mammalian expression vector.

For example, if pC4 (Accession No. 209646) is used, the human Fc portion can be ligated into the BamHI cloning site. Note that the 3' BamHI site should be destroyed. Next, the vector containing the human Fc portion is re-restricted with BamHI, linearizing the vector, and CRGCL polynucleotide, isolated by the PCR protocol described in Example 1, is ligated into this BamHI site.

Alternatively, a soluble CRGCL polypeptide, such as amino acids Met 1 to Lys 231, could also be fused to the Fc portion. For Example, a 5' primer: 5' CCGGTTAGATCTGCCATCATGGGGCGGCTGGTTCTG 3' (SEQ ID NO:28), having a Bgl II restriction site and a 3' primer: 5' GGCCGGTCTAGATTTGGACAGCTTTGGTTTG 3' (SEQ ID NO:29) could be used to PCR amino acids Met 1 to Lys 231. The amplified product could be fused to Fc to produce a Fc fusion protein, as set forth above, and ligated to the pC4 vector.

In either case, note that the polynucleotide is cloned without a stop codon, otherwise a fusion protein will not be produced. Moreover, if the naturally occurring signal sequence is used to produce the secreted protein, pC4 does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

The Fc fusions described above could also be inserted into the pA2 vector to express in Baculovirus systems, as set forth in Example 7, using techniques known in the art and described herein.

#### Human IgG Fc region:

CCGATCCGGAGCCCAAAATCTCTGACAAAACCTCAACATGCCCAACCGTGCCCAAGCACTGAATTGAG  
GGTGCACCGTCAGTCTTCTCTTCCCCCAAAACCAAGGACACCCCTCATGATCTCCCGGACTCTGA  
GGTCACATGCCGTGGTGGGACGTAAGCCACGAAGACCCCTGAGGTCAAGTTCAACTGGTACGTGGACG  
GCGTGGAGGTGCATTAATGCCAAGACAAAGCCGCGGAGGAGCACTACAACAGCACGTACCGTGTCTGTC

15

## 10

20

25

30

35

Such a method makes use of the fact that antibodies are themselves antigens, and therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the CRCGCL protein-specific antibody can be blocked by CRCGCL. Such antibodies comprise anti-idiotypic antibodies to the CRCGCL protein-specific antibody and can be used to immunize an animal to induce formation of further CRCGCL protein-specific antibodies.

It will be appreciated that Fab and F(ab')<sub>2</sub> and other fragments of the antibodies of the present invention may be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')<sub>2</sub> fragments). Alternatively, secreted CRCGCL protein-binding fragments can be produced through the application of recombinant DNA technology or through synthetic chemistry.

For in vivo use of antibodies in humans, it may be preferable to use "humanized" chimeric monoclonal antibodies. Such antibodies can be produced using genetic constructs derived from hybridoma cells producing the monoclonal antibodies described above. Methods for producing chimeric antibodies are known in the art. (See, for review, Morrison, Science 229:1202 (1985); Oi et al., BioTechniques 4:214 (1986); Cabilly et al., U.S. Patent No. 4,816,567; Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulianne et al., Nature 312:643 (1984); Neuberger et al., Nature 314:268 (1985).)

#### **Example 12: Production Of CRCGCL Protein For High-Throughput Screening Assays**

The following protocol produces a supernatant containing CRCGCL polypeptide to be tested. This supernatant can then be used in the Screening Assays described in Examples 14-21.

First, dilute Poly-D-Lysine (644 587 Boehringer-Mannheim) stock solution (1mg/ml in PBS) 1:20 in PBS (w/o calcium or magnesium 17-516F Biowhittaker) for a working solution of 50ug/ml. Add 200 ul of this solution to each well (24 well plates) and incubate at RT for 20 minutes. Be sure to distribute the solution over each well (note: a 12-channel pipetter may be used with tips on every other channel). Aspirate off the Poly-D-Lysine solution and rinse with 1ml PBS (Phosphate Buffered



Saline). The PBS should remain in the well until just prior to plating the cells and plates may be poly-lysine coated in advance for up to two weeks.

Plate 293T cells (do not carry cells past P+20) at  $2 \times 10^5$  cells/well in .5ml DMEM-Dulbecco's Modified Eagle Medium(with 4.5 G/L glucose and L-glutamine (12-604F Biowhittaker))/10% heat inactivated FBS(14-503F Biowhittaker)/1x Penstrep(17-602F Biowhittaker). Let the cells grow overnight.

The next day, mix together in a sterile solution basin: 300 ul Lipofectamine (18324-012 Gibco/BRL) and 5ml Optumem I (31985070 Gibco/BRL)/96-well plate. With a small volume multi-channel pipetter, aliquot approximately 2ug of an expression vector containing a polynucleotide insert, produced by the methods described in Examples 8-10, into an appropriately labeled 96-well round bottom plate. With a multi channel pipetter, add 50ul of the Lipofectamine/Optumem I mixture to each well. Pipette up and down gently to mix. Incubate at RT 15-45 minutes. After about 20 minutes, use a multi-channel pipetter to add 150ul Optumem I to each well. As a control, one plate of vector DNA lacking an insert should be transfected with each set of transfections.

Preferably, the transfection should be performed by tag-teaming the following tasks. By tag-teaming, hands on time is cut in half, and the cells do not spend too much time on PBS. First, person A aspirates off the media from four 24-well plates of cells, and then person B rinses each well with .5-1ml PBS. Person A then aspirates off PBS rinse, and person B, using a 12-channel pipetter with tips on every other channel, adds the 200ul of DNA/Lipofectamine/Optumem I complex to the odd wells first, then to the even wells, to each row on the 24-well plates. Incubate at 37 degree C for 6 hours.

While cells are incubating, prepare appropriate media, either 1%BSA in DMEM with 1x penstrep, or HGS CHO-5 media (116.6 mg/L of CaCl<sub>2</sub> (anhyd); 0.00130 mg/L CuSO<sub>4</sub>·5H<sub>2</sub>O; 0.050 mg/L of Fe(NO<sub>3</sub>)<sub>3</sub>·9H<sub>2</sub>O; 0.417 mg/L of FeSO<sub>4</sub>·7H<sub>2</sub>O; 311.80 mg/L of Kcl; 28.64 mg/L of MgCl<sub>2</sub>; 48.84 mg/L of MgSO<sub>4</sub>; 6995.50 mg/L of NaCl; 2400.0 mg/L of NaHCO<sub>3</sub>; 62.50 mg/L of NaH<sub>2</sub>PO<sub>4</sub>·H<sub>2</sub>O; 71.02 mg/L of Na<sub>2</sub>HPO<sub>4</sub>; .4320 mg/L of ZnSO<sub>4</sub>·7H<sub>2</sub>O; .002 mg/L of Arachidonic Acid ; 1.022 mg/L of Cholesterol; .070 mg/L of DL-alpha-Tocopherol-Acetate; 0.0520 mg/L of Linoleic Acid; 0.010 mg/L of Linolenic Acid; 0.010 mg/L of Myristic Acid; 0.010 mg/L of Oleic Acid; 0.010 mg/L of Palmitic Acid; 0.010 mg/L of Palmitic Acid; 100 mg/L of Pluronic F-68; 0.010 mg/L of Stearic Acid; 2.20 mg/L of Tween 80; 4551 mg/L of D-Glucose; 130.85 mg/ml of L- Alanine; 147.50 mg/ml of L-Arginine-HCL; 7.50 mg/ml of L-Asparagine-H<sub>2</sub>O; 6.65 mg/ml of L-Aspartic Acid;

29.56 mg/ml of L-Cystine-2HCL-H<sub>2</sub>O; 31.29 mg/ml of L-Cystine-2HCL; 7.35 mg/ml of L-Glutamic Acid; 365.0 mg/ml of L-Glutamine; 18.75 mg/ml of Glycine; 52.48 mg/ml of L-Histidine-HCL-H<sub>2</sub>O; 106.97 mg/ml of L-Isoleucine; 111.45 mg/ml of L-Leucine; 163.75 mg/ml of L-Lysine HCL; 32.34 mg/ml of L-Methionine; 68.48 mg/ml of L-Phenylalanine; 40.0 mg/ml of L-Proline; 26.25 mg/ml of L-Serine; 101.05 mg/ml of L-Threonine; 19.22 mg/ml of L-Tryptophan; 91.79 mg/ml of L-Tyrosine-2Na-2H<sub>2</sub>O; and 99.65 mg/ml of L-Valine; 0.0035 mg/L of Biotin; 3.24 mg/L of D-Ca Pantothenate; 11.78 mg/L of Choline Chloride; 4.65 mg/L of Folic Acid; 15.60 mg/L of D-Inositol; 3.02 mg/L of Niacinamide; 3.00 mg/L of Pyridoxal HCL; 0.031 mg/L of Pyridoxine HCL; 0.319 mg/L of Riboflavin; 3.17 mg/L of Thiamine HCL; 0.365 mg/L of Thymidine; 0.680 mg/L of Vitamin B<sub>12</sub>; 25 mM of HEPES Buffer; 2.39 mg/L of Na Hypoxanthine; 0.105 mg/L of Lipic Acid; 0.081 mg/L of Sodium Putrescine-2HCL; 55.0 mg/L of Sodium Pyruvate; 0.0067 mg/L of Sodium Selenite; 20uM of Ethanolamine; 0.122 mg/L of Ferric Citrate; 41.70 mg/L of Methyl-B-Cyclodextrin complexed with Linoleic Acid; 33.33 mg/L of Methyl-B-Cyclodextrin complexed with Oleic Acid; 10 mg/L of Methyl-B-Cyclodextrin complexed with Retinal Acetate. Adjust osmolarity to 327 mOsm) with 2mM glutamine and 1x penstrep. (BSA (81-068-3 Bayer) 100gm dissolved in 1L DMEM for a 10% BSA stock solution). Filter the media and collect 50 ul for endotoxin assay in 15ml polystyrene conical.

The transfection reaction is terminated, preferably by tag-teaming, at the end of the incubation period. Person A aspirates off the transfection media, while person B adds 1.5ml appropriate media to each well. Incubate at 37 degree C for 45 or 72 hours depending on the media used: 1%BSA for 45 hours or CHO-5 for 72 hours.

On day four, using a 300ul multichannel pipetter, aliquot 600ul in one 1ml deep well plate and the remaining supernatant into a 2ml deep well. The supernatants from each well can then be used in the assays described in Examples 14-21.

It is specifically understood that when activity is obtained in any of the assays described below using a supernatant, the activity originates from either the CRCGCL polypeptide directly (e.g., as a secreted protein) or by CRCGCL inducing expression of other proteins, which are then secreted into the supernatant. Thus, the invention further provides a method of identifying the protein in the supernatant characterized by an activity in a particular assay.

**Example 13: Construction of GAS Reporter Construct**

One signal transduction pathway involved in the differentiation and proliferation of cells is called the Jaks-STATs pathway. Activated proteins in the Jaks-STATs pathway bind to gamma activation site "GAS" elements or interferon sensitive responsive element ("ISRE"), located in the promoter of many genes. The binding of a protein to these elements alter the expression of the associated gene.

GAS and ISRE elements are recognized by a class of transcription factors called Signal Transducers and Activators of Transcription, or "STATs." There are six members of the STATs family. Stat1 and Stat3 are present in many cell types, as is Stat2 (as response to IFN-alpha is widespread). Stat4 is more restricted and is not in many cell types though it has been found in T helper class 1 cells after treatment with IL-12. Stat5 was originally called mammary growth factor, but has been found at higher concentrations in other cells including myeloid cells. It can be activated in tissue culture cells by many cytokines.

The STATs are activated to translocate from the cytoplasm to the nucleus upon tyrosine phosphorylation by a set of kinases known as the Janus Kinase ("Jaks") family. Jaks represent a distinct family of soluble tyrosine kinases and include Tyk2, Jak1, Jak2, and Jak3. These kinases display significant sequence similarity and are generally catalytically inactive in resting cells.

The Jaks are activated by a wide range of receptors summarized in the Table below. (Adapted from review by Schidler and Darnell, Ann. Rev. Biochem. 64:621-51 (1995).) A cytokine receptor family, capable of activating Jaks, is divided into two groups: (a) Class 1 includes receptors for IL-2, IL-3, IL-4, IL-6, IL-7, IL-9, IL-11, IL-12, IL-15, Epo, PRL, GH, G-CSF, GM-CSF, LIF, CNTF, and thrombopoietin; and (b) Class 2 includes IFN-a, IFN-g, and IL-10. The Class 1 receptors share a conserved cysteine motif (a set of four conserved cysteines and one tryptophan) and a WSXWS motif (a membrane proximal region encoding Trp-Ser-Xxx-Trp-Ser (SEQ ID NO:5)).

Thus, on binding of a ligand to a receptor, Jaks are activated, which in turn activate STATs, which then translocate and bind to GAS elements. This entire process is encompassed in the Jaks-STATs signal transduction pathway.

Therefore, activation of the Jaks-STATs pathway, reflected by the binding of the GAS or the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells. For example, growth factors and cytokines are known to activate the Jaks-STATs pathway. (See Table below.) Thus, by using GAS elements linked to reporter molecules, activators of the Jaks-STATs pathway can be identified.

There is preliminary data that CRGCL interacts with Jak1.

5	<u>or ISRE</u> <u>Ligand</u>	<u>JAKs</u>				<u>STATS</u>	<u>GAS(elements)</u>
		<u>tyk2</u>	<u>Jak1</u>	<u>Jak2</u>	<u>Jak3</u>		
	<u>IFN family</u>						
	IFN- $\alpha$ /B	+	+	-	-	1,2,3	ISRE
10	IFN-g (IRF1>Lys6>IFP)		+	+		1	GAS
	IL-10	+	?	?		1,3	
	<u>gp130 family</u>						
15	IL-6 (Pleiotrohic) (IRF1>Lys6>IFP)	+	+	+	?	1,3	GAS
	IL-11 (Pleiotrohic)	?	+	?	?	1,3	
	OnM (Pleiotrohic)	?	+	+	?	1,3	
	LIF (Pleiotrohic)	?	+	+	?	1,3	
20	CNTF (Pleiotrohic)	+/-	+	+	?	1,3	
	G-CSF (Pleiotrohic)	?	+	?	?	1,3	
	IL-12 (Pleiotrohic)	+	-	+	+	1,3	
	<u>g-C family</u>						
25	IL-2 (lymphocytes)	-	+	-	+	1,3,5	GAS
	IL-4 (lymph/myeloid)	-	+		+	6	GAS (IRF1 =
	IFP >> Ly6)(IgH)						
	IL-7 (lymphocytes)	-	+	-	+	5	GAS
	IL-9 (lymphocytes)	-	+	-	+	5	GAS
30	IL-13 (lymphocyte)	-	+	?	?	6	GAS
	IL-15	?	+	?	+	5	GAS
	<u>gp140 family</u>						
35	IL-3 (myeloid)	-	-	+	-	5	GAS
	(IRF1>IFP>>Ly6)						
	IL-5 (myeloid)	-	-	+	-	5	GAS
	GM-CSF (myeloid)	-	-	+	-	5	GAS
	<u>Growth hormone family</u>						
40	GH	?	-	+	-	5	
	PRL	?	+/-	+	-	1,3,5	
	EPO	?	-	+	-	5	GAS(B-
	CAS>IRF1=IFP>>Ly6)						
45	<u>Receptor Tyrosine Kinases</u>						
	EGF	?	+	+	-	1,3	GAS (IRF1)
	PDGF	?	+	+	-	1,3	
50	CSF-1	?	+	+	-	1,3	GAS (not IRF1)

To construct a synthetic GAS containing promoter element, which is used in the Biological Assays described in Examples 14-15, a PCR based strategy is employed to generate a GAS-SV40 promoter sequence. The 5' primer contains four tandem copies of the GAS binding site found in the IRF1 promoter and previously demonstrated to bind STATs upon induction with a range of cytokines (Rothman et al., Immunity 1:457-468 (1994)), although other GAS or ISRE elements can be used instead. The 5' primer also contains 18bp of sequence complementary to the SV40 early promoter sequence and is flanked with an XhoI site. The sequence of the 5' primer is:

5' GCGCCTCGAGATTTCCTCCGAAATCTAGATTTCCTCCGAAATGATTTCCTCCGAAATGATTTCCTCCGAAATATCTGCCATCTCAATTAG:3' (SEQ ID NO:6)

The downstream primer is complementary to the SV40 promoter and is flanked with a Hind III site: 5' GCGGCAAGCTTTTGGCAAAGCCTAGGC:3' (SEQ ID NO:7)

PCR amplification is performed using the SV40 promoter template present in the B-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI/Hind III and subcloned into B1.SK2- (Stratagene). Sequencing with forward and reverse primers confirms that the insert contains the following sequence:

5' CTCGAGATTTCCTCCGAAATCTAGATTTCCTCCGAAATGATTTCCTCCGAAATGATTTCCTCCGAAATATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCGCCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTATTTATGCAGAGGCCGAGGCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGGCAAAGCCTT:3' (SEQ ID NO:8)

With this GAS promoter element linked to the SV40 promoter, a GAS:SEAP2 reporter construct is next engineered. Here, the reporter molecule is a secreted alkaline phosphatase, or "SEAP." Clearly, however, any reporter molecule can be instead of SEAP, in this or in any of the other Examples. Well known reporter molecules that can be used instead of SEAP include chloramphenicol acetyltransferase (CAT), luciferase, alkaline phosphatase, B-galactosidase, green fluorescent protein (GFP), or any protein detectable by an antibody.

The above sequence confirmed synthetic GAS-SV40 promoter element is subcloned into the pSEAP-Promoter vector obtained from Clontech using HindIII and XhoI, effectively replacing the SV40 promoter with the amplified GAS:SV40 promoter element, to create the GAS-SEAP vector. However, this vector does not

contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

Thus, in order to generate mammalian stable cell lines expressing the GAS-SEAP reporter, the GAS-SEAP cassette is removed from the GAS-SEAP vector using  
5   Sall and NotI, and inserted into a backbone vector containing the neomycin resistance gene, such as pGFP-1 (Clontech), using these restriction sites in the multiple cloning site, to create the GAS-SEAP/Neo vector. Once this vector is transfected into mammalian cells, this vector can then be used as a reporter molecule for GAS binding as described in Examples 14-15.

10   Other constructs can be made using the above description and replacing GAS with a different promoter sequence. For example, construction of reporter molecules containing NFK-B and EGR promoter sequences are described in Examples 16 and 17. However, many other promoters can be substituted using the protocols described in these Examples. For instance, SRE, IL-2, NFAT, or Osteocalcin promoters can  
15   be substituted, alone or in combination (e.g., GAS/NF-KB/EGR, GAS/NF-KB, IL-2/NFAT, or NF-KB/GAS). Similarly, other cell lines can be used to test reporter construct activity, such as HELA (epithelial), HUVEC (endothelial), Reh (B-cell), Saos-2 (osteoblast), HUVAC (aortic), or Cardiomyocyte.

#### 20   Example 14: High-Throughput Screening Assay for T-cell Activity.

The following protocol is used to assess T-cell activity of CRGCL by determining whether CRGCL supernatant proliferates and/or differentiates T-cells. T-cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 13. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-  
25   STATS signal transduction pathway. The T-cell used in this assay is Jurkat T-cells (ATCC Accession No. TIB-152), although Molt-3 cells (ATCC Accession No. CRL-1552) and Molt-4 cells (ATCC Accession No. CRL-1582) cells can also be used.

Jurkat T-cells are lymphoblastic CD4+ Th1 helper cells. In order to generate stable cell lines, approximately 2 million Jurkat cells are transfected with the GAS-  
30   SEAP/neo vector using DMRIE-C (Life Technologies)(transfection procedure described below). The transfected cells are seeded to a density of approximately 20,000 cells per well and transfectants resistant to 1 mg/ml gentamicin selected. Resistant colonies are expanded and then tested for their response to increasing concentrations of interferon gamma. The dose response of a selected clone is  
35   demonstrated.

Specifically, the following protocol will yield sufficient cells for 75 wells containing 200 ul of cells. Thus, it is either scaled up, or performed in multiple to

generate sufficient cells for multiple 96 well plates. Jurkat cells are maintained in RPMI + 10% serum with 1% Pen-Strep. Combine 2.5 ml of OPTI-MEM (Life Technologies) with 10 ug of plasmid DNA in a T25 flask. Add 2.5 ml OPTI-MEM containing 50 ul of DMRIE-C and incubate at room temperature for 15-45 mins.

- 5        During the incubation period, count cell concentration, spin down the required number of cells ( $10^7$  per transfection), and resuspend in OPTI-MEM to a final concentration of  $10^7$  cells/ml. Then add 1ml of  $1 \times 10^7$  cells in OPTI-MEM to T25 flask and incubate at 37 degree C for 6 hrs. After the incubation, add 10 ml of RPMI + 15% serum.

- 10       The Jurkat:GAS-SEAP stable reporter lines are maintained in RPMI + 10% serum, 1 mg/ml Genticin, and 1% Pen-Strep. These cells are treated with supernatants containing CRGCL polypeptides or CRGCL induced polypeptides as produced by the protocol described in Example 12.

- 15       On the day of treatment with the supernatant, the cells should be washed and resuspended in fresh RPMI + 10% serum to a density of 500,000 cells per ml. The exact number of cells required will depend on the number of supernatants being screened. For one 96 well plate, approximately 10 million cells (for 10 plates, 100 million cells) are required.

- 20       Transfer the cells to a triangular reservoir boat, in order to dispense the cells into a 96 well dish, using a 12 channel pipette. Using a 12 channel pipette, transfer 200 ul of cells into each well (therefore adding 100,000 cells per well).

- 25       After all the plates have been seeded, 50 ul of the supernatants are transferred directly from the 96 well plate containing the supernatants into each well using a 12 channel pipette. In addition, a dose of exogenous interferon gamma (0.1, 1.0, 10 ng) is added to wells H9, H10, and H11 to serve as additional positive controls for the assay.

- 30       The 96 well dishes containing Jurkat cells treated with supernatants are placed in an incubator for 48 hrs (note: this time is variable between 48-72 hrs). 35 ul samples from each well are then transferred to an opaque 96 well plate using a 12 channel pipette. The opaque plates should be covered (using sellophene covers) and stored at -20 degree C until SEAP assays are performed according to Example 18. The plates containing the remaining treated cells are placed at 4 degree C and serve as a source of material for repeating the assay on a specific well if desired.

- 35       As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate Jurkat T cells. Over 30 fold induction is typically observed in the positive control wells.

**Example 15: High-Throughput Screening Assay Identifying Myeloid Activity**

The following protocol is used to assess myeloid activity of CRGCL by determining whether CRGCL proliferates and/or differentiates myeloid cells. Myeloid cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 13. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks/STATS signal transduction pathway. The myeloid cell used in this assay is U937, a pre-monocyte cell line, although TF-1, HL60, or KG1 can be used.

To transiently transfect U937 cells with the GAS/SEAP/Neo construct produced in Example 13, a DEAE-Dextran method (Kharbanda et. al., 1994, Cell Growth & Differentiation, 5:259-265) is used. First, harvest  $2 \times 10^7$  U937 cells and wash with PBS. The U937 cells are usually grown in RPMI 1640 medium containing 10% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 mg/ml streptomycin.

Next, suspend the cells in 1 ml of 20 mM Tris-HCl (pH 7.4) buffer containing 0.5 mg/ml DEAE-Dextran, 8  $\mu$ g GAS-SEAP2 plasmid DNA, 140 mM NaCl, 5 mM KCl, 375  $\mu$ M  $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$ , 1 mM  $\text{MgCl}_2$ , and 675  $\mu$ M  $\text{CaCl}_2$ . Incubate at 37 degree C for 45 min.

Wash the cells with RPMI 1640 medium containing 10% FBS and then resuspend in 10 ml complete medium and incubate at 37 degree C for 36 hr.

The GAS-SEAP/U937 stable cells are obtained by growing the cells in 400  $\mu$ g/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 400  $\mu$ g/ml G418 for couple of passages.

These cells are tested by harvesting  $1 \times 10^6$  cells (this is enough for ten 96-well plates assay) and wash with PBS. Suspend the cells in 200  $\mu$ l above described growth medium, with a final density of  $5 \times 10^5$  cells/ml. Plate 200  $\mu$ l cells per well in the 96-well plate (or  $1 \times 10^3$  cells/well).

Add 50  $\mu$ l of the supernatant prepared by the protocol described in Example 12. Incubate at 37 degree C for 48 to 72 hr. As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate U937 cells. Over 30 fold induction is typically observed in the positive control wells. SEAP assay the supernatant according to the protocol described in Example 18.



**Example 16: High-Throughput Screening Assay Identifying Neuronal Activity.**

When cells undergo differentiation and proliferation, a group of genes are activated through many different signal transduction pathways. One of these genes, EGR1 (early growth response gene 1), is induced in various tissues and cell types upon activation. The promoter of EGR1 is responsible for such induction. Using the EGR1 promoter linked to reporter molecules, activation of cells can be assessed by CRCGCL.

Particularly, the following protocol is used to assess neuronal activity in PC12 cell lines. PC12 cells (rat pheochromocytoma cells) are known to proliferate and/or differentiate by activation with a number of mitogens, such as TPA (tetradecanoyl phorbol acetate), NGF (nerve growth factor), and EGF (epidermal growth factor). The EGR1 gene expression is activated during this treatment. Thus, by stably transfecting PC12 cells with a construct containing an EGR promoter linked to SEAP reporter, activation of PC12 cells by CRCGCL can be assessed.

The EGR/SEAP reporter construct can be assembled by the following protocol. The EGR-1 promoter sequence (-633 to +1) (Sakamoto K et al., Oncogene 6:867-871 (1991)) can be PCR amplified from human genomic DNA using the following primers:

5' GCGCTCGAGGGATGACAGCGATAGAACCCCGG -3' (SEQ ID NO:9)

5' GCGAAGCTTCGCGACTCCCCGGATCCGCCTC-3' (SEQ ID NO:10)

Using the GAS:SEAP/Neo vector produced in Example 13, EGR1 amplified product can then be inserted into this vector. Linearize the GAS:SEAP/Neo vector using restriction enzymes XhoI/HindIII, removing the GAS/SV40 stuffer. Restrict the EGR1 amplified product with these same enzymes. Ligate the vector and the EGR1 promoter.

To prepare 96 well-plates for cell culture, two mls of a coating solution (1:30 dilution of collagen type I (Upstate Biotech Inc. Cat#08-115) in 30% ethanol (filter sterilized)) is added per one 10 cm plate or 50 ml per well of the 96-well plate, and allowed to air dry for 2 hr.

PC12 cells are routinely grown in RPMI-1640 medium (Bio Whittaker) containing 10% horse serum (JRH BIOSCIENCES, Cat. # 12449-78P), 5% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 ug/ml streptomycin on a precoated 10 cm tissue culture dish. One to four split is done every three to four days. Cells are removed from the plates by scraping and resuspended with pipetting up and down for more than 15 times.

Transtect the EGR/SEAP/Neo construct into PC12 using the Lipofectamine protocol described in Example 12. EGR-SEAP/PC12 stable cells are obtained by growing the cells in 300 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 300 ug/ml G418 for couple of passages.

To assay for neuronal activity, a 10 cm plate with cells around 70 to 80% confluent is screened by removing the old medium. Wash the cells once with PBS (Phosphate buffered saline). Then starve the cells in low serum medium (RPMI-1640 containing 1% horse serum and 0.5% FBS with antibiotics) overnight.

The next morning, remove the medium and wash the cells with PBS. Scrape off the cells from the plate, suspend the cells well in 2 ml low serum medium. Count the cell number and add more low serum medium to reach final cell density as  $5 \times 10^5$  cells/ml.

Add 200 ul of the cell suspension to each well of 96-well plate (equivalent to  $1 \times 10^5$  cells/well). Add 50 ul supernatant produced by Example 12, 37 degree C for 48 to 72 hr. As a positive control, a growth factor known to activate PC12 cells through EGR can be used, such as 50 ng/ul of Neuronal Growth Factor (NGF). Over fifty-fold induction of SEAP is typically seen in the positive control wells. SEAP assay the supernatant according to Example 18.

#### **Example 17: High-Throughput Screening Assay for T-cell Activity**

NF-KB (Nuclear Factor KB) is a transcription factor activated by a wide variety of agents including the inflammatory cytokines IL-1 and TNF, CD30 and CD40, lymphotoxin-alpha and lymphotoxin-beta, by exposure to LPS or thrombin, and by expression of certain viral gene products. As a transcription factor, NF-KB regulates the expression of genes involved in immune cell activation, control of apoptosis (NF- KB appears to shield cells from apoptosis), B and T-cell development, anti-viral and antimicrobial responses, and multiple stress responses.

In non-stimulated conditions, NF- KB is retained in the cytoplasm with I-KB (Inhibitor KB). However, upon stimulation, I- KB is phosphorylated and degraded, causing NF- KB to shuttle to the nucleus, thereby activating transcription of target genes. Target genes activated by NF- KB include IL-2, IL-6, GM-CSF, ICAM-1 and class I MHC.

Due to its central role and ability to respond to a range of stimuli, reporter constructs utilizing the NF-KB promoter element are used to screen the supernatants produced in Example 12. Activators or inhibitors of NF-KB would be useful in

treating diseases. For example, inhibitors of NF-KB could be used to treat those diseases related to the acute or chronic activation of NF-KB, such as rheumatoid arthritis.

To construct a vector containing the NF-KB promoter element, a PCR based strategy is employed. The upstream primer contains four tandem copies of the NF-KB binding site (GGGGACTTTCCC) (SEQ ID NO:11), 18 bp of sequence complementary to the 5' end of the SV40 early promoter sequence, and is flanked with an XhoI site:

5' GCGGCCTCGAGGGGACTTTCCCGGGGACTTTCCGGGGACTTTCCGGGA  
CTTTCCATCCTGCCATCTCAATTAG:3' (SEQ ID NO:12)

The downstream primer is complementary to the 3' end of the SV40 promoter and is flanked with a Hind III site:

5' GCGGCAAGCTTTTTGCAAAGCCTAGGC:3' (SEQ ID NO:7)

PCR amplification is performed using the SV40 promoter template present in the pB-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI and Hind III and subcloned into BLSK2- (Stratagene). Sequencing with the T7 and T3 primers confirms the insert contains the following sequence:

5' CTCGAGGGGACTTTCCCGGGGACTTTCCGGGGACTTTCCGGGACTTTCC  
ATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCCTAACTCCGCCC  
ATCCCGCCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGA  
CTAATTTTTTTTTATTTATGCAGAGGCCGAGGCCGCCCTCGGCCTCTGAGCTA  
TTCCAGAAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAA  
GCTT:3' (SEQ ID NO:13)

Next, replace the SV40 minimal promoter element present in the pSEAP2-promoter plasmid (Clontech) with this NF-KB/SV40 fragment using XhoI and HindIII. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

In order to generate stable mammalian cell lines, the NF-KB/SV40/SEAP cassette is removed from the above NF-KB/SEAP vector using restriction enzymes SalI and NotI, and inserted into a vector containing neomycin resistance. Particularly, the NF-KB/SV40/SEAP cassette was inserted into pGFP-1 (Clontech), replacing the GFP gene, after restricting pGFP-1 with SalI and NotI.

Once NF-KB/SV40/SEAP/Neo vector is created, stable Jurkat T-cells are created and maintained according to the protocol described in Example 14. Similarly, the method for assaying supernatants with these stable Jurkat T-cells is also described

in Example 14. As a positive control, exogenous TNF alpha (0.1, 1, 10 ng) is added to wells H9, H10, and H11, with a 5-10 fold activation typically observed.

#### **Example 18: Assay for SEAP Activity**

- 5 As a reporter molecule for the assays described in Examples 14-17, SEAP activity is assayed using the Tropix Phospho-light Kit (Cat. BP-400) according to the following general procedure. The Tropix Phospho-light Kit supplies the Dilution, Assay, and Reaction Buffers used below.

- 10 Prime a dispenser with the 2.5x Dilution Buffer and dispense 15 ul of 2.5x dilution buffer into Optiplates containing 35 ul of a supernatant. Seal the plates with a plastic sealer and incubate at 65 degree C for 30 min. Separate the Optiplates to avoid uneven heating.

- 15 Cool the samples to room temperature for 15 minutes. Empty the dispenser and prime with the Assay Buffer. Add 50 ml Assay Buffer and incubate at room temperature 5 min. Empty the dispenser and prime with the Reaction Buffer (see the table below). Add 50 ul Reaction Buffer and incubate at room temperature for 20 minutes. Since the intensity of the chemiluminescent signal is time dependent, and it takes about 10 minutes to read 5 plates on luminometer, one should treat 5 plates at each time and start the second set 10 minutes later.

- 20 Read the relative light unit in the luminometer. Set H12 as blank, and print the results. An increase in chemiluminescence indicates reporter activity.

#### **Reaction Buffer Formulation:**

# of plates	Rxn buffer diluent (ml)	CSPD (ml)
10	60	3
11	65	3.25
12	70	3.5
13	75	3.75
14	80	4
15	85	4.25
16	90	4.5
17	95	4.75
18	100	5
19	105	5.25
20	110	5.5
21	115	5.75
22	120	6
23	125	6.25
24	130	6.5
25	135	6.75
26	140	7
27	145	7.25
28	150	7.5

29	155	7.75
30	160	8
31	165	8.25
32	170	8.5
33	175	8.75
34	180	9
35	185	9.25
36	190	9.5
37	195	9.75
38	200	10
39	205	10.25
40	210	10.5
41	215	10.75
42	220	11
43	225	11.25
44	230	11.5
45	235	11.75
46	240	12
47	245	12.25
48	250	12.5
49	255	12.75
50	260	13

**Example 19: High-Throughput Screening Assay Identifying Changes in Small Molecule Concentration and Membrane Permeability**

Binding of a ligand to a receptor is known to alter intracellular levels of small molecules, such as calcium, potassium, sodium, and pH, as well as alter membrane potential. These alterations can be measured in an assay to identify supernatants which bind to receptors of a particular cell. Although the following protocol describes an assay for calcium, this protocol can easily be modified to detect changes in potassium, sodium, pH, membrane potential, or any other small molecule which is detectable by a fluorescent probe.

The following assay uses Fluorometric Imaging Plate Reader ("FLIPR") to measure changes in fluorescent molecules (Molecular Probes) that bind small molecules. Clearly, any fluorescent molecule detecting a small molecule can be used instead of the calcium fluorescent molecule, fluo-3, used here.

For adherent cells, seed the cells at 10,000 -20,000 cells/well in a Co-star black 96-well plate with clear bottom. The plate is incubated in a CO<sub>2</sub> incubator for 20 hours. The adherent cells are washed two times in Biotek washer with 200 ul of HBSS (Hank's Balanced Salt Solution) leaving 100 ul of buffer after the final wash.

A stock solution of 1 mg/ml fluo-3 is made in 10% pluronic acid DMSO. To load the cells with fluo-3, 50 ul of 12 ug/ml fluo-3 is added to each well. The plate is incubated at 37 degree C in a CO<sub>2</sub> incubator for 60 min. The plate is washed four times in the Biotek washer with HBSS leaving 100 ul of buffer.

For non-adherent cells, the cells are spun down from culture media. Cells are re-suspended to  $2-5 \times 10^5$  cells/ml with HBSS in a 50-ml conical tube. 4  $\mu$ l of 1 mg/ml fluo-3 solution in 10% pluronic acid DMSO is added to each ml of cell suspension. The tube is then placed in a 37 degree C water bath for 30-60 min. The cells are washed twice with HBSS, resuspended to  $1 \times 10^6$  cells/ml, and dispensed into a microplate, 100  $\mu$ l/well. The plate is centrifuged at 1000 rpm for 5 min. The plate is then washed once in Denley CellWash with 200  $\mu$ l, followed by an aspiration step to 100  $\mu$ l final volume.

For a non-cell based assay, each well contains a fluorescent molecule, such as fluo-3. The supernatant is added to the well, and a change in fluorescence is detected.

To measure the fluorescence of intracellular calcium, the FLIPR is set for the following parameters: (1) System gain is 300-800 mW; (2) Exposure time is 0.4 second; (3) Camera F-stop is F/2; (4) Excitation is 488 nm; (5) Emission is 530 nm; and (6) Sample addition is 50  $\mu$ l. Increased emission at 530 nm indicates an extracellular signaling event caused by the a molecule, either CRGCL or a molecule induced by CRGCL, which has resulted in an increase in the intracellular  $Ca^{++}$  concentration.

#### **Example 20: High-Throughput Screening Assay Identifying Tyrosine Kinase Activity**

The Protein Tyrosine Kinases (PTK) represent a diverse group of transmembrane and cytoplasmic kinases. Within the Receptor Protein Tyrosine Kinase (RPTK) group are receptors for a range of mitogenic and metabolic growth factors including the PDGF, FGF, EGF, NGF, HGF and Insulin receptor subfamilies. In addition there are a large family of RPTKs for which the corresponding ligand is unknown. Ligands for RPTKs include mainly secreted small proteins, but also membrane-bound and extracellular matrix proteins.

Activation of RPTK by ligands involves ligand-mediated receptor dimerization, resulting in transphosphorylation of the receptor subunits and activation of the cytoplasmic tyrosine kinases. The cytoplasmic tyrosine kinases include receptor associated tyrosine kinases of the src-family (e.g., src, yes, lck, lyn, fyn) and non-receptor linked and cytosolic protein tyrosine kinases, such as the Jak family, members of which mediate signal transduction triggered by the cytokine superfamily of receptors (e.g., the Interleukins, Interferons, GM-CSF, and Leptin).

Because of the wide range of known factors capable of stimulating tyrosine kinase activity, identifying whether CRGCL or a molecule induced by CRGCL is

capable of activating tyrosine kinase signal transduction pathways is of interest. Therefore, the following protocol is designed to identify such molecules capable of activating the tyrosine kinase signal transduction pathways.

Seed target cells (e.g., primary keratinocytes) at a density of approximately 25,000 cells per well in a 96 well Loprodyne Silent Screen Plates purchased from Nalge Nunc (Naperville, IL). The plates are sterilized with two 30 minute rinses with 100% ethanol, rinsed with water and dried overnight. Some plates are coated for 2 hr with 100  $\mu$ l of cell culture grade type I collagen (50 mg/ml), gelatin (2%) or polylysine (50 mg/ml), all of which can be purchased from Sigma Chemicals (St. Louis, MO) or 10% Matrigel purchased from Becton Dickinson (Bedford, MA), or calf serum, rinsed with PBS and stored at 4 degree C. Cell growth on these plates is assayed by seeding 5,000 cells/well in growth medium and indirect quantitation of cell number through use of alamarBlue as described by the manufacturer Alamar Biosciences, Inc. (Sacramento, CA) after 48 hr. Falcon plate covers #3071 from Becton Dickinson (Bedford, MA) are used to cover the Loprodyne Silent Screen Plates. Falcon Microtest III cell culture plates can also be used in some proliferation experiments.

To prepare extracts, A431 cells are seeded onto the nylon membranes of Loprodyne plates (20,000/200 $\mu$ l/well) and cultured overnight in complete medium. Cells are quiesced by incubation in serum-free basal medium for 24 hr. After 5-20 minutes treatment with EGF (60ng/ml) or 50  $\mu$ l of the supernatant produced in Example 12, the medium was removed and 100  $\mu$ l of extraction buffer ((20 mM HEPES pH 7.5, 0.15 M NaCl, 1% Triton X-100, 0.1% SDS, 2 mM Na<sub>3</sub>VO<sub>4</sub>, 2 mM Na<sub>4</sub>P<sub>2</sub>O<sub>7</sub> and a cocktail of protease inhibitors (# 1836170) obtained from Boehringer Mannheim (Indianapolis, IN) is added to each well and the plate is shaken on a rotating shaker for 5 minutes at 4°C. The plate is then placed in a vacuum transfer manifold and the extract filtered through the 0.45  $\mu$ m membrane bottoms of each well using house vacuum. Extracts are collected in a 96-well catch/assay plate in the bottom of the vacuum manifold and immediately placed on ice. To obtain extracts clarified by centrifugation, the content of each well, after detergent solubilization for 5 minutes, is removed and centrifuged for 15 minutes at 4 degree C at 16,000 x g.

Test the filtered extracts for levels of tyrosine kinase activity. Although many methods of detecting tyrosine kinase activity are known, one method is described here

Generally, the tyrosine kinase activity of a supernatant is evaluated by determining its ability to phosphorylate a tyrosine residue on a specific substrate (a biotinylated peptide). Biotinylated peptides that can be used for this purpose include PSK1 (corresponding to amino acids 6-20 of the cell division kinase cdc2-p34) and PSK2 (corresponding to amino acids 1-17 of gastrin). Both peptides are substrates for a range of tyrosine kinases and are available from Boehringer Mannheim.

The tyrosine kinase reaction is set up by adding the following components in order. First, add 10ul of 5uM Biotinylated Peptide, then 10ul ATP/Mg<sup>2+</sup> (5mM ATP/50mM MgCl<sub>2</sub>), then 10ul of 8x Assay Buffer (40mM imidazole hydrochloride, pH7.3, 40 mM beta-glycerophosphate, 1mM EGTA, 100mM MgCl<sub>2</sub>, 5 mM MnCl<sub>2</sub>, 0.5 mg/ml BSA), then 5ul of Sodium Vanadate(1mM), and then 5ul of water. Mix the components gently and preincubate the reaction mix at 30 degree C for 2 min. Initiate the reaction by adding 10ul of the control enzyme or the filtered supernatant.

The tyrosine kinase assay reaction is then terminated by adding 10 ul of 120mM EDTA and place the reactions on ice.

Tyrosine kinase activity is determined by transferring 50 ul aliquot of reaction mixture to a microtiter plate (MTP) module and incubating at 37 degree C for 20 min. This allows the streptavidin coated 96 well plate to associate with the biotinylated peptide. Wash the MTP module with 300ul/well of PBS four times. Next add 75 ul of anti-phosphotyrosine antibody conjugated to horse radish peroxidase(anti-P-Tyr-POD(0.5u/ml)) to each well and incubate at 37 degree C for one hour. Wash the well as above.

Next add 100ul of peroxidase substrate solution (Boehringer Mannheim) and incubate at room temperature for at least 5 mins (up to 30 min). Measure the absorbance of the sample at 405 nm by using ELISA reader. The level of bound peroxidase activity is quantitated using an ELISA reader and reflects the level of tyrosine kinase activity.

#### **Example 21: High-Throughput Screening Assay Identifying Phosphorylation Activity**

As a potential alternative and/or complement to the assay of protein tyrosine kinase activity described in Example 20, an assay which detects activation (phosphorylation) of major intracellular signal transduction intermediates can also be used. For example, as described below one particular assay can detect tyrosine phosphorylation of the Erk-1 and Erk-2 kinases. However, phosphorylation of other molecules, such as Raf, JNK, p38 MAP, Map kinase kinase (MEK), MEK kinase,



Src, Muscle specific kinase (MuSK), IRAK, Tec, and Janus, as well as any other phosphoserine, phosphotyrosine, or phosphothreonine molecule, can be detected by substituting these molecules for Erk-1 or Erk-2 in the following assay.

Specifically, assay plates are made by coating the wells of a 96-well ELISA plate with 0.1ml of protein G (1ug/ml) for 2 hr at room temp. (RT). The plates are then rinsed with PBS and blocked with 3% BSA/PBS for 1 hr at RT. The protein G plates are then treated with 2 commercial monoclonal antibodies (100ng/well) against Erk-1 and Erk-2 (1 hr at RT) (Santa Cruz Biotechnology). (To detect other molecules, this step can easily be modified by substituting a monoclonal antibody detecting any of the above described molecules.) After 3-5 rinses with PBS, the plates are stored at 4 degree C until use.

A431 cells are seeded at 20,000/well in a 96-well Loprodyne filterplate and cultured overnight in growth medium. The cells are then starved for 48 hr in basal medium (DMEM) and then treated with EGF (6ng/well) or 50 ul of the supernatants obtained in Example 12 for 5-20 minutes. The cells are then solubilized and extracts filtered directly into the assay plate.

After incubation with the extract for 1 hr at RT, the wells are again rinsed. As a positive control, a commercial preparation of MAP kinase (10ng/well) is used in place of A431 extract. Plates are then treated with a commercial polyclonal (rabbit) antibody (1ug/ml) which specifically recognizes the phosphorylated epitope of the Erk-1 and Erk-2 kinases (1 hr at RT). This antibody is biotinylated by standard procedures. The bound polyclonal antibody is then quantitated by successive incubations with Europium-streptavidin and Europium fluorescence enhancing reagent in the Wallac DELFIA instrument (time-resolved fluorescence). An increased fluorescent signal over background indicates a phosphorylation by CRCGCL or a molecule induced by CRCGCL.

#### **Example 22: Method of Determining Alterations in the CRCGCL Gene**

RNA isolated from entire families or individual patients presenting with a phenotype of interest (such as a disease) is be isolated. cDNA is then generated from these RNA samples using protocols known in the art. (See, Sambrook.) The cDNA is then used as a template for PCR, employing primers surrounding regions of interest in SEQ ID NO:1. Suggested PCR conditions consist of 35 cycles at 95 degree C for 30 seconds; 60-120 seconds at 52-58 degree C; and 60-120 seconds at 70 degree C, using buffer solutions described in Sidransky, D., et al., Science 252:706 (1991).

PCR products are then sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTherm Polymerase (Epicentre Technologies). The intron-exon borders of selected exons of CRCGCL is also determined and genomic PCR products analyzed to confirm the results. PCR products harboring suspected mutations in CRCGCL is then cloned and sequenced to validate the results of the direct sequencing.

PCR products of CRCGCL are cloned into T-tailed vectors as described in Holton, T.A. and Graham, M.W., Nucleic Acids Research, 19:1156 (1991) and sequenced with T7 polymerase (United States Biochemical). Affected individuals are identified by mutations in CRCGCL not present in unaffected individuals.

Genomic rearrangements are also observed as a method of determining alterations in the CRCGCL gene. Genomic clones isolated according to Example 2 are nick-translated with digoxigenindeoxy-uridine 5'-triphosphate (Boehringer Mannheim), and FISH performed as described in Johnson, Cg. et al., Methods Cell Biol. 35:73-99 (1991). Hybridization with the labeled probe is carried out using a vast excess of human cot-1 DNA for specific hybridization to the CRCGCL genomic locus.

Chromosomes are counterstained with 4,6-diamino-2-phenylidole and propidium iodide, producing a combination of C- and R-bands. Aligned images for precise mapping are obtained using a triple-band filter set (Chroma Technology, Brattleboro, VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters. (Johnson, Cv. et al., Genet. Anal. Tech. Appl., 8:75 (1991).) Image collection, analysis and chromosomal fractional length measurements are performed using the ISee Graphical Program System. (Inovision Corporation, Durham, NC.) Chromosome alterations of the genomic region of CRCGCL (hybridized by the probe) are identified as insertions, deletions, and translocations. These CRCGCL alterations are used as a diagnostic marker for an associated disease.

### **Example 23: Method of Detecting Abnormal Levels of CRCGCL in a Biological Sample**

CRCGCL polypeptides can be detected in a biological sample, and if an increased or decreased level of CRCGCL is detected, this polypeptide is a marker for a particular phenotype. Methods of detection are numerous, and thus, it is understood that one skilled in the art can modify the following assay to fit their particular needs.

For example, antibody-sandwich ELISAs are used to detect CRCGCL in a sample, preferably a biological sample. Wells of a microtiter plate are coated with

specific antibodies to CRCGCL, at a final concentration of 0.2 to 10 ug/ml. The antibodies are either monoclonal or polyclonal and are produced by the method described in Example 11. The wells are blocked so that non-specific binding of CRCGCL to the well is reduced.

5       The coated wells are then incubated for > 2 hours at RT with a sample containing CRCGCL. Preferably, serial dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled water to remove unbounded CRCGCL.

10       Next, 50 ul of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbounded conjugate.

15       Add 75 ul of 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution to each well and incubate 1 hour at room temperature. Measure the reaction by a microtiter plate reader. Prepare a standard curve, using serial dilutions of a control sample, and plot CRCGCL polypeptide concentration on the X-axis (log scale) and fluorescence or absorbance of the Y-axis (linear scale). Interpolate the concentration of the CRCGCL in the sample using the standard curve.

20

#### **Example 24: Formulating a Polypeptide**

25       The CRCGCL composition will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the CRCGCL polypeptide alone), the site of delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

30       As a general proposition, the total pharmaceutically effective amount of CRCGCL administered parenterally per dose will be in the range of about 1ug/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, CRCGCL is typically administered at a dose rate of about 1 ug/kg/hour to about 50 ug/kg/hour, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag  
35       solution may also be employed. The length of treatment needed to observe changes

and the interval following treatment for responses to occur appears to vary depending on the desired effect.

Pharmaceutical compositions containing CRCGCL are administered orally, rectally, parenterally, intracisternally, intravaginally, intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), buccally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

CRCGCL is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semi-permeable polymer matrices in the form of shaped articles, e.g., films, or microcapsules. Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919; EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U. et al., *Biopolymers* 22:547-556 (1983)), poly (2-hydroxyethyl methacrylate) (R. Langer et al., *J. Biomed. Mater. Res.* 15:167-277 (1981), and R. Langer, *Chem. Tech.* 12:98-105 (1982)), ethylene vinyl acetate (R. Langer et al.) or poly-D-(-)-3-hydroxybutyric acid (EP 133,988). Sustained-release compositions also include liposomally entrapped CRCGCL polypeptides. Liposomes containing the CRCGCL are prepared by methods known per se: DE 3,218,121; Epstein et al., *Proc. Natl. Acad. Sci. USA* 82:3688-3692 (1985); Hwang et al., *Proc. Natl. Acad. Sci. USA* 77:4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal secreted polypeptide therapy.

For parenteral administration, in one embodiment, CRCGCL is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to polypeptides.

Generally, the formulations are prepared by contacting CRCGCL uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is

a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

5       The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues)  
10 polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrans; chelating agents such as EDTA; sugar  
15 alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

CRCGCL is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers  
20 will result in the formation of polypeptide salts.

CRCGCL used for therapeutic administration can be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial  
25 having a stopper pierceable by a hypodermic injection needle.

CRCGCL polypeptides ordinarily will be stored in unit or multi-dose containers, for example, sealed ampoules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1% (w/v) aqueous  
30 CRCGCL polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized CRCGCL polypeptide using bacteriostatic Water-for-Injection.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical  
35 compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency

of manufacture, use or sale for human administration. In addition, CRCGCL may be employed in conjunction with other therapeutic compounds.

**Example 25: Method of Treating Decreased Levels of CRCGCL**

5       The present invention relates to a method for treating an individual in need of a decreased level of CRCGCL activity in the body comprising administering to such an individual a composition comprising a therapeutically effective amount of CRCGCL antagonist. Preferred antagonists for use in the present invention are CRCGCL-specific antibodies.

10       Moreover, it will be appreciated that conditions caused by a decrease in the standard or normal expression level of CRCGCL in an individual can be treated by administering CRCGCL, preferably in the secreted form. Thus, the invention also provides a method of treatment of an individual in need of an increased level of CRCGCL polypeptide comprising administering to such an individual a  
15       pharmaceutical composition comprising an amount of CRCGCL to increase the activity level of CRCGCL in such an individual.

For example, a patient with decreased levels of CRCGCL polypeptide receives a daily dose 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the polypeptide is in the secreted form. The exact details of the dosing scheme, based  
20       on administration and formulation, are provided in Example 24.

**Example 26: Method of Treating Increased Levels of CRCGCL**

25       The present invention also relates to a method for treating an individual in need of an increased level of CRCGCL activity in the body comprising administering to such an individual a composition comprising a therapeutically effective amount of CRCGCL or an agonist thereof.

Antisense technology is used to inhibit production of CRCGCL. This technology is one example of a method of decreasing levels of CRCGCL polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer.

30       For example, a patient diagnosed with abnormally increased levels of CRCGCL is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period if the treatment was well tolerated. The formulation of the antisense polynucleotide is provided in Example 24.

35

**Example 27: Method of Treatment Using Gene Therapy - Ex Vivo**

One method of gene therapy transplants fibroblasts, which are capable of expressing CRCGCL polypeptides, onto a patient. Generally, fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of the flask and fresh media (e.g., Ham's F12 media, with 10% FBS, penicillin and streptomycin) is added. The flasks are then incubated at 37 degree C for approximately one week.

At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks.

pMV-7 (Kirschmeier, P.T. et al., DNA, 7:219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

The cDNA encoding CRCGCL can be amplified using PCR primers which correspond to the 5' and 3' end sequences respectively as set forth in Example 1. Preferably, the 5' primer contains an EcoRI site and the 3' primer includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The ligation mixture is then used to transform bacteria HB101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector contains properly inserted CRCGCL.

The amphotropic pA317 or GP+am12 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the CRCGCL gene is then added to the media and the packaging cells transduced with the vector. The packaging cells now produce infectious viral particles containing the CRCGCL gene (the packaging cells are now referred to as producer cells).

Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media, containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is

removed from a sub-confluent plate of fibroblasts and quickly replaced with the media from the producer cells. This media is removed and replaced with fresh media. If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether CRCGCL protein is produced.

The engineered fibroblasts are then transplanted onto the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads.

10

#### **Example 28: Gene Therapy Using Endogenous CRCGCL Gene**

Another method of gene therapy according to the present invention involves operably associating the endogenous CRCGCL sequence with a promoter via homologous recombination as described, for example, in U.S. Patent No. 5,641,670, issued June 24, 1997; International Publication No. WO 96/29411, published September 26, 1996; International Publication No. WO 94/12650, published August 4, 1994; Koller *et al.*, *Proc. Natl. Acad. Sci. USA* 86:8932-8935 (1989); and Zijlstra *et al.*, *Nature* 342:435-438 (1989). This method involves the activation of a gene which is present in the target cells, but which is not expressed in the cells, or is expressed at a lower level than desired.

Polynucleotide constructs are made which contain a promoter and targeting sequences, which are homologous to the 5' non-coding sequence of endogenous CRCGCL, flanking the promoter. The targeting sequence will be sufficiently near the 5' end of CRCGCL so the promoter will be operably linked to the endogenous sequence upon homologous recombination. The promoter and the targeting sequences can be amplified using PCR. Preferably, the amplified promoter contains distinct restriction enzyme sites on the 5' and 3' ends. Preferably, the 3' end of the first targeting sequence contains the same restriction enzyme site as the 5' end of the amplified promoter and the 5' end of the second targeting sequence contains the same restriction site as the 3' end of the amplified promoter.

The amplified promoter and the amplified targeting sequences are digested with the appropriate restriction enzymes and subsequently treated with calf intestinal phosphatase. The digested promoter and digested targeting sequences are added together in the presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The construct is size fractionated on an agarose gel then purified by phenol extraction and ethanol precipitation.



In this Example, the polynucleotide constructs are administered as naked polynucleotides via electroporation. However, the polynucleotide constructs may also be administered with transfection-facilitating agents, such as liposomes, viral sequences, viral particles, precipitating agents, etc. Such methods of delivery are known in the art.

Once the cells are transfected, homologous recombination will take place which results in the promoter being operably linked to the endogenous CRCGCL sequence. This results in the expression of CRCGCL in the cell. Expression may be detected by immunological staining, or any other method known in the art.

Fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in DMEM + 10% fetal calf serum. Exponentially growing or early stationary phase fibroblasts are trypsinized and rinsed from the plastic surface with nutrient medium. An aliquot of the cell suspension is removed for counting, and the remaining cells are subjected to centrifugation. The supernatant is aspirated and the pellet is resuspended in 5 ml of electroporation buffer (20 mM HEPES pH 7.3, 137 mM NaCl, 5 mM KCl, 0.7 mM  $\text{Na}_2\text{HPO}_4$ , 6 mM dextrose). The cells are recentrifuged, the supernatant aspirated, and the cells resuspended in electroporation buffer containing 1 mg/ml acetylated bovine serum albumin. The final cell suspension contains approximately  $3 \times 10^6$  cells/ml. Electroporation should be performed immediately following resuspension.

Plasmid DNA is prepared according to standard techniques. For example, to construct a plasmid for targeting to the CRCGCL locus, plasmid pUC18 (MBI Fermentas, Amherst, NY) is digested with HindIII. The CMV promoter is amplified by PCR with an XbaI site on the 5' end and a BamHI site on the 3' end. Two CRCGCL non-coding sequences are amplified via PCR: one CRCGCL non-coding sequence (CRCGCL fragment 1) is amplified with a HindIII site at the 5' end and an Xba site at the 3' end; the other CRCGCL non-coding sequence (CRCGCL fragment 2) is amplified with a BamHI site at the 5' end and a HindIII site at the 3' end. The CMV promoter and CRCGCL fragments are digested with the appropriate enzymes (CMV promoter - XbaI and BamHI; CRCGCL fragment 1 - XbaI; CRCGCL fragment 2 - BamHI) and ligated together. The resulting ligation product is digested with HindIII, and ligated with the HindIII-digested pUC18 plasmid.

Plasmid DNA is added to a sterile cuvette with a 0.4 cm electrode gap (Bio-Rad). The final DNA concentration is generally at least 120  $\mu\text{g/ml}$ . 0.5 ml of the cell suspension (containing approximately  $1.5 \times 10^6$  cells) is then added to the cuvette, and the cell suspension and DNA solutions are gently mixed. Electroporation is performed with a Gene-Pulser apparatus (Bio-Rad). Capacitance and voltage are set at

960  $\mu$ F and 250-300 V, respectively. As voltage increases, cell survival decreases, but the percentage of surviving cells that stably incorporate the introduced DNA into their genome increases dramatically. Given these parameters, a pulse time of approximately 14-20 mSec should be observed.

5 Electroporated cells are maintained at room temperature for approximately 5 min, and the contents of the cuvette are then gently removed with a sterile transfer pipette. The cells are added directly to 10 ml of prewarmed nutrient media (DMEM with 15% calf serum) in a 10 cm dish and incubated at 37 degree C. The following day, the media is aspirated and replaced with 10 ml of fresh media and incubated for a  
10 further 16-24 hours.

The engineered fibroblasts are then injected into the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads. The fibroblasts now produce the protein product. The fibroblasts can then be introduced into a patient as described above.

15

#### **Example 29: Method of Treatment Using Gene Therapy - In Vivo**

Another aspect of the present invention is using *in vivo* gene therapy methods to treat disorders, diseases and conditions. The gene therapy method relates to the introduction of naked nucleic acid (DNA, RNA, and antisense DNA or RNA)  
20 CRGCL sequences into an animal to increase or decrease the expression of the CRGCL polypeptide. The CRGCL polynucleotide may be operatively linked to a promoter or any other genetic elements necessary for the expression of the CRGCL polypeptide by the target tissue. Such gene therapy and delivery techniques and methods are known in the art, see, for example, WO90/11092, WO98/11779; U.S.  
25 Patent NO. 5693622, 5705151, 5580859; Tabata H. et al. (1997) Cardiovasc. Res. 35(3):470-479. Chao J et al. (1997) Pharmacol. Res. 35(6):517-522, Wolff J.A. (1997) Neuromuscul. Disord. 7(5):314-318, Schwartz B. et al. (1996) Gene Ther. 3(5):405-411, Tsurumi Y. et al. (1996) Circulation 94(12):3281-3290 (incorporated herein by reference).

30 The CRGCL polynucleotide constructs may be delivered by any method that delivers injectable materials to the cells of an animal, such as, injection into the interstitial space of tissues (heart, muscle, skin, lung, liver, intestine and the like). The CRGCL polynucleotide constructs can be delivered in a pharmaceutically acceptable liquid or aqueous carrier.

35 The term "naked" polynucleotide, DNA or RNA, refers to sequences that are free from any delivery vehicle that acts to assist, promote, or facilitate entry into the cell, including viral sequences, viral particles, liposome formulations, lipofectin or

precipitating agents and the like. However, the CRCGCL polynucleotides may also be delivered in liposome formulations (such as those taught in Feigner P.L. et al. (1995) Ann. NY Acad. Sci. 772:126-139 and Abdallah B. et al. (1995) Biol. Cell 85(1):1-7) which can be prepared by methods well known to those skilled in the art.

5       The CRCGCL polynucleotide vector constructs used in the gene therapy method are preferably constructs that will not integrate into the host genome nor will they contain sequences that allow for replication. Any strong promoter known to those skilled in the art can be used for driving the expression of DNA. Unlike other gene therapies techniques, one major advantage of introducing naked nucleic acid  
10 sequences into target cells is the transitory nature of the polynucleotide synthesis in the cells. Studies have shown that non-replicating DNA sequences can be introduced into cells to provide production of the desired polypeptide for periods of up to six months.

      The CRCGCL polynucleotide construct can be delivered to the interstitial space of tissues within the an animal, including of muscle, skin, brain, lung, liver,  
15 spleen, bone marrow, thymus, heart, lymph, blood, bone, cartilage, pancreas, kidney, gall bladder, stomach, intestine, testis, ovary, uterus, rectum, nervous system, eye, gland, and connective tissue. Interstitial space of the tissues comprises the intercellular fluid, mucopolysaccharide matrix among the reticular fibers of organ tissues, elastic fibers in the walls of vessels or chambers, collagen fibers of fibrous  
20 tissues, or that same matrix within connective tissue ensheathing muscle cells or in the lacunae of bone. It is similarly the space occupied by the plasma of the circulation and the lymph fluid of the lymphatic channels. Delivery to the interstitial space of muscle tissue is preferred for the reasons discussed below. They may be conveniently delivered by injection into the tissues comprising these cells. They are preferably  
25 delivered to and expressed in persistent, non-dividing cells which are differentiated, although delivery and expression may be achieved in non-differentiated or less completely differentiated cells, such as, for example, stem cells of blood or skin fibroblasts. *In vivo* muscle cells are particularly competent in their ability to take up and express polynucleotides.

30       For the naked CRCGCL polynucleotide injection, an effective dosage amount of DNA or RNA will be in the range of from about 0.05 g/kg body weight to about 50 mg/kg body weight. Preferably the dosage will be from about 0.005 mg/kg to about 20 mg/kg and more preferably from about 0.05 mg/kg to about 5 mg/kg. Of course, as the artisan of ordinary skill will appreciate, this dosage will vary according to the  
35 tissue site of injection. The appropriate and effective dosage of nucleic acid sequence can readily be determined by those of ordinary skill in the art and may depend on the condition being treated and the route of administration. The preferred route of

administration is by the parenteral route of injection into the interstitial space of tissues. However, other parenteral routes may also be used, such as, inhalation of an aerosol formulation particularly for delivery to lungs or bronchial tissues, throat or mucous membranes of the nose. In addition, naked CRCGCL polynucleotide constructs can be delivered to arteries during angioplasty by the catheter used in the procedure

The dose response effects of injected CRCGCL polynucleotide in muscle *in vivo* is determined as follows. Suitable CRCGCL template DNA for production of mRNA coding for CRCGCL polypeptide is prepared in accordance with a standard recombinant DNA methodology. The template DNA, which may be either circular or linear, is either used as naked DNA or complexed with liposomes. The quadriceps muscles of mice are then injected with various amounts of the template DNA.

Five to six week old female and male Balb/C mice are anesthetized by intraperitoneal injection with 0.3 ml of 2.5% Avertin. A 1.5 cm incision is made on the anterior thigh, and the quadriceps muscle is directly visualized. The CRCGCL template DNA is injected in 0.1 ml of carrier in a 1 cc syringe through a 27 gauge needle over one minute, approximately 0.5 cm from the distal insertion site of the muscle into the knee and about 0.2 cm deep. A suture is placed over the injection site for future localization, and the skin is closed with stainless steel clips.

After an appropriate incubation time (e.g., 7 days) muscle extracts are prepared by excising the entire quadriceps. Every fifth 15 um cross-section of the individual quadriceps muscles is histochemically stained for CRCGCL protein expression. A time course for CRCGCL protein expression may be done in a similar fashion except that quadriceps from different mice are harvested at different times. Persistence of CRCGCL DNA in muscle following injection may be determined by Southern blot analysis after preparing total cellular DNA and HIRT supernatants from injected and control mice. The results of the above experimentation in mice can be used to extrapolate proper dosages and other treatment parameters in humans and other animals using CRCGCL naked DNA.

30

#### **Example 30: CRCGCL Transgenic Animals.**

The CRCGCL polypeptides can also be expressed in transgenic animals. Animals of any species, including, but not limited to, mice, rats, rabbits, hamsters, guinea pigs, pigs, micro-pigs, goats, sheep, cows and non-human primates, e.g., baboons, monkeys, and chimpanzees may be used to generate transgenic animals. In a specific embodiment, techniques described herein or otherwise known in the art, are

35

used to express polypeptides of the invention in humans, as part of a gene therapy protocol.

Any technique known in the art may be used to introduce the transgene (i.e., polynucleotides of the invention) into animals to produce the founder lines of transgenic animals. Such techniques include, but are not limited to, pronuclear microinjection (Paterson et al., Appl. Microbiol. Biotechnol. 40:691-698 (1994); Carver et al., Biotechnology (NY) 11:1263-1270 (1993); Wright et al., Biotechnology (NY) 9:830-834 (1991); and Hoppe et al., U.S. Pat. No. 4,873,191 (1989)); retrovirus mediated gene transfer into germ lines (Van der Putten et al., Proc. Natl. Acad. Sci., USA 82:6148-6152 (1985)); blastocysts or embryos; gene targeting in embryonic stem cells (Thompson et al., Cell 56:313-321 (1989)); electroporation of cells or embryos (Lo, 1983, Mol Cell. Biol. 3:1803-1814 (1983)); introduction of the polynucleotides of the invention using a gene gun (see, e.g., Ulmer et al., Science 259:1745 (1993)); introducing nucleic acid constructs into embryonic pluripotent stem cells and transferring the stem cells back into the blastocyst; and sperm-mediated gene transfer (Lavitrano et al., Cell 57:717-723 (1989)); etc. For a review of such techniques, see Gordon, "Transgenic Animals," Intl. Rev. Cytol. 115:171-229 (1989), which is incorporated by reference herein in its entirety.

Any technique known in the art may be used to produce transgenic clones containing polynucleotides of the invention, for example, nuclear transfer into enucleated oocytes of nuclei from cultured embryonic, fetal, or adult cells induced to quiescence (Campbell et al., Nature 380:64-66 (1996); Wilmut et al., Nature 385:810-813 (1997)).

The present invention provides for transgenic animals that carry the transgene in all their cells, as well as animals which carry the transgene in some, but not all their cells, i.e., mosaic animals or chimeric. The transgene may be integrated as a single transgene or as multiple copies such as in concatamers, e.g., head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and activated in a particular cell type by following, for example, the teaching of Lasko et al. (Lasko et al., Proc. Natl. Acad. Sci. USA 89:6232-6236 (1992)). The regulatory sequences required for such a cell-type specific activation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art. When it is desired that the polynucleotide transgene be integrated into the chromosomal site of the endogenous gene, gene targeting is preferred.

Briefly, when such a technique is to be utilized, vectors containing some nucleotide sequences homologous to the endogenous gene are designed for the purpose of integrating, via homologous recombination with chromosomal sequences,

into and disrupting the function of the nucleotide sequence of the endogenous gene. The transgene may also be selectively introduced into a particular cell type, thus inactivating the endogenous gene in only that cell type, by following, for example, the teaching of Gu et al. (Gu et al., Science 265:103-106 (1994)). The regulatory  
5 sequences required for such a cell-type specific inactivation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

Once transgenic animals have been generated, the expression of the recombinant gene may be assayed utilizing standard techniques. Initial screening may be accomplished by Southern blot analysis or PCR techniques to analyze animal  
10 tissues to verify that integration of the transgene has taken place. The level of mRNA expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques which include, but are not limited to, Northern blot analysis of tissue samples obtained from the animal, *in situ* hybridization analysis, and reverse transcriptase-PCR (rt-PCR). Samples of transgenic gene-expressing tissue may also  
15 be evaluated immunocytochemically or immunohistochemically using antibodies specific for the transgene product.

Once the founder animals are produced, they may be bred, inbred, outbred, or crossbred to produce colonies of the particular animal. Examples of such breeding strategies include, but are not limited to: outbreeding of founder animals with more  
20 than one integration site in order to establish separate lines; inbreeding of separate lines in order to produce compound transgenics that express the transgene at higher levels because of the effects of additive expression of each transgene; crossing of heterozygous transgenic animals to produce animals homozygous for a given integration site in order to both augment expression and eliminate the need for  
25 screening of animals by DNA analysis; crossing of separate homozygous lines to produce compound heterozygous or homozygous lines; and breeding to place the transgene on a distinct background that is appropriate for an experimental model of interest.

Transgenic animals of the invention have uses which include, but are not  
30 limited to, animal model systems useful in elaborating the biological function of CRCGCL polypeptides, studying conditions and/or disorders associated with aberrant CRCGCL expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

### 35 **Example 31: CRCGCL Knock-Out Animals.**

Endogenous CRCGCL gene expression can also be reduced by inactivating or "knocking out" the CRCGCL gene and/or its promoter using targeted homologous

recombination. (E.g., see Smithies et al., Nature 317:230-234 (1985); Thomas & Capecchi, Cell 51:503-512 (1987); Thompson et al., Cell 5:313-321 (1989); each of which is incorporated by reference herein in its entirety). For example, a mutant, non-functional polynucleotide of the invention (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous polynucleotide sequence (either the coding regions or regulatory regions of the gene) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells that express polypeptides of the invention *in vivo*. In another embodiment, techniques known in the art are used to generate knockouts in cells that contain, but do not express the gene of interest. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the targeted gene. Such approaches are particularly suited in research and agricultural fields where modifications to embryonic stem cells can be used to generate animal offspring with an inactive targeted gene (e.g., see Thomas & Capecchi 1987 and Thompson 1989, *supra*). However this approach can be routinely adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site *in vivo* using appropriate viral vectors that will be apparent to those of skill in the art.

In further embodiments of the invention, cells that are genetically engineered to express the polypeptides of the invention, or alternatively, that are genetically engineered not to express the polypeptides of the invention (e.g., knockouts) are administered to a patient *in vivo*. Such cells may be obtained from the patient (i.e., animal, including human) or an MHC compatible donor and can include, but are not limited to fibroblasts, bone marrow cells, blood cells (e.g., lymphocytes), adipocytes, muscle cells, endothelial cells etc. The cells are genetically engineered *in vitro* using recombinant DNA techniques to introduce the coding sequence of polypeptides of the invention into the cells, or alternatively, to disrupt the coding sequence and/or endogenous regulatory sequence associated with the polypeptides of the invention, e.g., by transduction (using viral vectors, and preferably vectors that integrate the transgene into the cell genome) or transfection procedures, including, but not limited to, the use of plasmids, cosmids, YACs, naked DNA, electroporation, liposomes, etc. The coding sequence of the polypeptides of the invention can be placed under the control of a strong constitutive or inducible promoter or promoter/enhancer to achieve expression, and preferably secretion, of the CRGCL polypeptides. The engineered cells which express and preferably secrete the polypeptides of the invention can be introduced into the patient systemically, e.g., in the circulation, or intraperitoneally.

Alternatively, the cells can be incorporated into a matrix and implanted in the body, e.g., genetically engineered fibroblasts can be implanted as part of a skin graft.

genetically engineered endothelial cells can be implanted as part of a lymphatic or vascular graft. (See, for example, Anderson et al. U.S. Patent No. 5,399,349, and Mulligan & Wilson, U.S. Patent No. 5,460,959 each of which is incorporated by reference herein in its entirety).

5        When the cells to be administered are non-autologous or non-MHC compatible cells, they can be administered using well known techniques which prevent the development of a host immune response against the introduced cells. For example, the cells may be introduced in an encapsulated form which, while allowing for an exchange of components with the immediate extracellular environment, does not allow  
10       the introduced cells to be recognized by the host immune system.

Knock-out animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of CRCGCL polypeptides, studying conditions and/or disorders associated with aberrant CRCGCL expression, and in screening for compounds effective in ameliorating such  
15       conditions and/or disorders. For example, a knock-out mouse can be made using the sequences disclosed as AA008694 and W98372, herein incorporated by reference in their entirety.

It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples. Numerous  
20       modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

The entire disclosure of each document cited (including patents, patent applications, journal articles, abstracts, laboratory manuals, books, or other disclosures) in the Background of the Invention, Detailed Description, and Examples  
25       is hereby incorporated herein by reference. Moreover, the sequence listing from the priority documents are herein incorporated by reference.



Res	Pos	Germ... Alpha	Chou... Alpha	Germ... Beta	Chou... Beta	Germ... Turn	Chou... Turn	Germ... Coil	Kyle... Hydro	Hopp... Hydro	Eisen... Alpha	Eisen... Beta	Karpl... Flex	Jamas... Antig	Emmi... Surfa
Met	1			B					-0.56	0.01	*	*		-0.10	0.59
Gly	2			B					-0.98	0.23	*	*		-0.10	0.34
Arg	3		A	B					-1.40	0.49	*	*		-0.60	0.22
Leu	4		A	B					-1.30	0.74	*	*		-0.60	0.18
Val	5		A	B					-1.26	1.04	*	*		-0.60	0.20
Leu	6		A	B					-1.24	1.04	*	*		-0.60	0.10
Leu	7	A	A						-1.49	1.54	*	*		-0.60	0.12
Trp	8	A	A						-2.46	1.36	*	*		-0.60	0.17
Gly	9	A	A						-2.34	1.36	*	*		-0.60	0.15
Ala	10	A	A						-2.30	1.46	*	*		-0.60	0.16
Ala	11	A	A						-2.30	1.46	*	*		-0.60	0.12
Val	12		A	B					-1.83	1.23	*	*		-0.60	0.10
Phe	13		A	B					-1.89	1.23	*	*		-0.60	0.10
Leu	14		A	B					-1.83	1.16	*	*		-0.60	0.10
Leu	15	A					T	C	-1.84	1.57				-0.20	0.14
Gly	16						T		-1.84	1.54				0.00	0.16
Gly	17					T	T		-1.80	1.26				0.20	0.19
Trp	18	A					T		-1.44	1.26				-0.20	0.19
Met	19			B					-0.63	1.00	*	*		-0.40	0.19
Ala	20			B					-0.17	0.97				-0.40	0.34
Leu	21			B					-0.17	0.97				-0.40	0.32
Gly	22						T	C	-0.41	0.49			F	0.32	0.32
Gln	23						T	C	-0.71	0.37			F	0.79	0.32
Gly	24						T	C	-0.11	0.37			F	0.96	0.39
Gly	25						T	C	0.13	-0.31			F	1.73	0.69
Ala	26							C	0.09	-0.31			F	1.70	0.39
Ala	27	A							0.43	-0.07	*	*	F	1.33	0.29
Glu	28	A			B				-0.46	-0.10	*	*		0.81	0.52
Gly	29	A			B				-0.11	0.16	*	*		0.04	0.36
Val	30	A			B				-0.66	0.06	*	*		0.13	0.61
Gln	31			B	B				-0.96	0.24	*	*		-0.30	0.25
Ile	32			B	B				-0.61	0.93	*	*		-0.60	0.18
Gln	33			B	B				-1.31	1.26	*	*		-0.60	0.37
Ile	34			B	B				-0.97	1.40	*	*		-0.60	0.19
Ile	35			B	B				-0.92	1.40	*	*		-0.60	0.43
Tyr	36			B	B				-0.92	1.40	*	*		-0.60	0.20
Phe	37			B	B				-0.34	1.00	*	*		-0.60	0.50
Asn	38				B			C	-1.20	0.80	*	*		-0.25	1.03
Leu	39				B			C	-0.31	0.76	*	*		-0.40	0.49
Glu	40			B	B				-0.28	0.40	*	*		-0.60	0.98
Thr	41			B	B				-0.34	0.26	*	*		-0.30	0.45
Val	42			B	B				0.07	0.34	*	*		-0.30	0.79
Gln	43			B	B				0.07	0.57	*	*		-0.60	0.48
Val	44			B	B				0.29	0.97	*	*		-0.60	0.53
Thr	45			B	B				-0.01	0.99	*	*		-0.60	0.73
Trp	46			B	B				0.34	0.73	*	*		-0.60	0.56
Asn	47	A			B				0.96	0.33	*	*		-0.15	1.51
Ala	48				B				0.56	0.44	*	*	F	0.44	1.64
Ser	49						T	C	1.62	0.34	*	*	F	1.28	2.10
Lys	50						T		1.62	-0.57	*	*	F	2.72	2.55
Tyr	51						T		1.91	-0.49	*	*	F	2.76	3.55
Ser	52						T		1.10	-0.59	*	*	F	3.40	4.38
Arg	53						T		1.38	-0.29	*	*	F	2.36	1.80
Thr	54			B	B				0.98	0.20	*	*	F	1.02	1.66
Asn	55			B	B				0.90	0.23	*	*	F	0.68	1.07
Leu	56			B	B				0.90	0.34	*	*		0.04	0.75
Thr	57			B	B				1.31	1.10	*	*		-0.60	0.81
Phe	58			B	B				0.50	0.61	*	*		-0.60	0.99
His	59			B	B				0.81	1.00	*	*		-0.45	1.34
Tyr	60			B	B				0.47	0.71	*	*		-0.45	1.15
Arg	61			B	B				1.28	0.66	*	*		-0.45	1.32
Phe	62				B		T		1.59	-0.13	*	*		1.19	1.62
Asn	63						T		1.70	-0.63	*	*	F	2.38	1.79
Gly	64						T		1.49	-0.89	*	*	F	2.57	0.92
Asp	65						T		1.73	-0.13	*	*	F	2.75	1.67
Glu	66						T		1.62	-0.91	*	*	F	3.40	1.73
Ala	67						T		1.56	-0.91	*	*	F	2.86	3.04
Tyr	68						T		1.34	-0.77	*	*		2.22	0.97
Asp	69						T		1.69	-0.29	*	*	F	1.73	0.81
Gln	70						T		1.44	0.11	*	*	F	0.94	1.29
Cys	71			B			T		0.63	0.37	*	*	F	0.40	1.29
Thr	72			B			T		0.41	0.30	*	*	F	0.25	0.64
Asn	73			B			T		0.66	0.99	*	*		-0.20	0.30
Tyr	74			B					0.66	0.99	*	*		-0.20	0.98
Leu	75			B					0.31	0.41	*	*		-0.25	1.18
Leu	76			B					0.94	0.36	*	*	F	-0.10	0.73
Gln	77			B					0.94	0.46	*	*	F	0.00	0.63

Res	Pos	Gami Alpha	Chou Alpha	Gami Beta	Chou Beta	Gami Turn	Chou Turn	Gami Coil	Kyte Hydro	Hopp Hydro	Eisen Alpha	Eisen Beta	Karpi Flex	James Anig	Emmi Sund
Glu	78			B					0.64	0.19			F	0.70	1.12
Gly	79					T			0.54	0.11			F	1.95	1.73
His	80					T			0.69	-0.37			F	2.20	1.02
Thr	81					T	T		0.59	0.20			F	2.50	0.32
Ser	82					T	T		-0.12	0.49			F	1.35	0.25
Gly	83					T	T		-0.12	0.74			F	1.13	0.16
Cys	84			B			T		-0.37	0.24					0.60
Leu	85			B			T		-0.33	0.26					-0.05
Leu	86	A	A						-0.02	0.13					0.30
Asp	87	A	A						0.39	-0.16					0.30
Ala	88	A	A						0.73	-0.73					0.90
Glu	89	A	A						1.40	-1.41			F	0.90	1.67
Gln	90	A	A				T		1.32	-2.10			F	1.30	3.79
Arg	91	A	A				T		1.32	1.41			F	1.30	3.79
Asp	92	A					T		1.08	-1.23			F	1.30	2.63
Asp	93	A					T		0.97	-0.47			F	1.00	1.25
Ile	94			B	B		T		0.67	0.09					1.13
Leu	95			B	B				-0.22	0.30					0.30
Tyr	96			B	B				0.22	0.99					-0.33
Phe	97			B	B				-0.22	0.99					0.40
Ser	98			B	B				-0.57	0.70					-0.60
Ile	99			B	B				0.01	0.44					0.17
Arg	100						T		0.79	0.17					0.47
Asn	101					T	T		0.82	-0.11			F	0.35	0.92
Gly	102					T	T		0.57	-0.07			F	1.49	0.58
Thr	103				B				0.27	-0.11			F	2.37	0.97
His	104				B			C	0.27	-0.11			F	2.80	0.38
Pro	105			B	B			C	0.84	0.67			F	1.77	2.17
Val	106			B	B				0.14	0.76			F	0.59	0.82
Phe	107			B	B				-0.16	0.83			F	-0.04	0.44
Thr	108			B	B				0.30	0.73			F	-0.32	0.55
Ala	109			B	B				0.32	0.23			F	-0.60	0.45
Ser	110	A		B	B				-0.24	0.71			F	-0.30	0.56
Arg	111	A		B	B				-0.89	0.69			F	-0.60	0.80
Tyr	112			B	B				-0.28	0.54			F	-0.80	0.31
Met	113			B	B				0.18	0.81			F	-0.60	0.47
Val	114			B	B				-0.32	1.07			F	-0.60	0.73
Tyr	115			B	B				0.31	1.37			F	-0.60	0.85
Tyr	116			B	B				0.40	1.37			F	-0.60	0.36
Leu	117			B	B				-0.01	0.89			F	-0.60	0.58
Lys	118								-0.02	0.66			F	-0.25	1.07
Pro	119							C	0.37	0.40			F	0.05	1.32
Ser	120					T	T	C	1.27	0.07			F	0.90	1.65
Ser	121					T	T	C	1.48	-0.69			F	1.70	1.62
Pro	122			B	B			C	0.87	-0.87			F	2.70	3.94
Lys	123			B	B				1.79	-0.23			F	3.00	2.68
His	124			B	B	T			1.04	-0.68			F	2.00	1.29
Val	125			B	B				0.96	0.26			F	2.40	1.88
Arg	126			B	B				0.97	-0.26			F	1.25	1.21
Phe	127			B	B				1.23	0.23			F	0.95	1.06
Ser	128			B	B				1.44	0.73			F	-0.10	0.55
Tyr	129					T			1.40	0.63			F	-0.38	0.55
His	130					T			0.84	-0.01			F	0.19	1.29
Gln	131				B	T		C	0.84	0.49			F	1.11	1.10
Asp	132			B	B	T			0.42	0.34			F	0.03	1.28
Ala	133			B	B				0.27	0.44			F	0.20	0.71
Val	134			B	B				0.26	0.17			F	-0.12	0.98
Thr	135			B	B				-0.12	0.16			F	-0.24	0.53
Val	136			B	B				-0.39	0.33			F	-0.26	0.44
Thr	137			B	B				-0.39	0.71			F	-0.28	0.14
Cys	138			B	B				-1.20	0.21			F	0.60	0.19
Ser	139			B	B		T		-0.91	0.26			F	-0.30	0.43
Asp	140			B	B		T		-0.30	0.16			F	0.10	0.24
Leu	141			B	B		T		-0.33	0.27			F	0.47	0.44
Ser	142					T	T		0.52	0.21			F	0.69	0.48
Tyr	143					T	T		0.02	-0.36			F	0.76	0.88
Gly	144					T	T		-0.12	-0.06			F	2.13	1.10
Asp	145					T	T		-0.07	0.63			F	2.20	0.54
Leu	146	A		B	B				-0.07	0.70			F	1.08	0.54
Leu	147	A		B	B				-0.11	0.31			F	0.46	0.63
Tyr	148	A		B	B				0.19	0.20			F	0.14	0.70
Glu	149	A		B	B				0.19	0.17			F	-0.08	0.52
Val	150			B	B				0.64	0.93			F	-0.30	0.54
Gln	151			B	B				0.34	0.24			F	-0.45	1.03
Tyr	152					T			0.94	-0.06			F	0.15	2.45
Arg	153			B					1.06	-0.39			F	0.45	2.10
Ser	154						T	C	1.30	0.40			F	1.39	1.87
									0.99	-0.24			F	0.58	2.19
													F	2.22	2.11

Res	Pos	Gamm Alpha	Chou Alpha	Gamm Beta	Chou Beta	Gamm Turn	Chou Turn	Gamm Coi	Kyle- Hydro	Hopp- Hydro	Eisen Alpha	Eisen Beta	Karpi- Flex	James- Arlig	Emm Surfa
Pro	155						T	C	1.24	0.16			F	2.50	1.94
Phe	156					T	T	C	1.56	-0.91			F	3.40	1.72
Asp	157						T	C	1.80	0.00			F	1.95	1.35
Thr	158							C	1.39	0.01			F	1.75	1.51
Glu	159	A							1.73	0.03			F	2.12	2.33
Trp	160	A					T		1.94	-0.81			F	2.65	2.80
Gln	161	A					T		2.64	-0.41			F	2.36	3.36
Ser	162					T	T		2.64	-0.90			F	3.40	3.36
Lys	163					T	T		2.64	-0.50			F	2.76	5.13
Gln	164					T	T		1.98	-0.93			F	2.69	4.28
Glu	165					T	T		2.27	-0.76			F	2.52	1.71
Asn	166					T	T		1.41	-0.74			F	2.55	1.38
Thr	167					T	T		1.10	0.10			F	1.93	0.59
Cys	168			B			T		0.47	-0.01			F	1.70	0.49
Asn	169			B			T		0.47	0.67				0.48	0.21
Val	170				B				0.12	0.27				0.21	0.26
Thr	171			B	B				-0.59	0.21				0.04	0.47
Ile	172		A	B	B				-0.38	0.33				-0.13	0.24
Glu	173	A	A		B				-0.30	-0.07				0.30	0.55
Gly	174	A	A						-0.30	-0.21				0.45	0.38
Leu	175	A	A						0.60	-0.70			F	0.80	0.95
Asp	176	A	A						0.24	-1.39			F	0.90	1.09
Ala	177	A	A						0.89	-0.81			F	0.75	0.59
Glu	178	A	A						0.59	-0.49			F	0.80	1.13
Lys	179	A	A		B				0.23	-0.79				0.50	0.90
Cys	180	A	A		B				0.76	0.00				-0.30	0.77
Tyr	181	A	A		B				-0.10	0.41				-0.60	0.47
Ser	182	A	A		B				0.60	1.06				-0.60	0.17
Phe	183	A	A		B				-0.26	1.06				-0.60	0.64
Trp	184	A	A		B				-0.26	1.13				-0.60	0.30
Val	185	A	A		B				-0.18	0.37				-0.30	0.45
Arg	186	A	A		B				-0.53	0.49				-0.60	0.53
Val	187	A	A		B				-0.23	0.31				-0.30	0.49
Lys	188	A	A		B				0.47	-0.60				0.75	1.15
Ala	189		A						-0.10	-1.24				0.60	0.98
Met	190	A	A						0.51	-0.60				0.60	0.98
Glu	191		A	B					0.06	-0.49				0.30	0.77
Asp	192		A	B					0.70	-0.06				0.30	0.76
Val	193		A	B					0.66	-0.13				0.45	1.18
Tyr	194		A	B					0.93	-0.74				0.75	1.14
Gly	195			B					1.29	-0.26			F	0.85	0.98
Pro	196					T	T		1.08	0.50			F	0.50	2.08
Asp	197					T	T		0.78	0.29			F	0.80	2.05
Thr	198						T	C	1.63	0.09			F	1.48	2.77
Tyr	199			B					1.59	-0.51			F	1.66	3.00
Pro	200						T	C	1.63	-0.03			F	2.04	1.89
Ser	201					T	T		1.84	0.36			F	1.92	1.75
Asp	202					T	T		0.99	-0.13			F	2.80	1.94
Trp	203					T	T		0.99	-0.24			F	2.37	0.93
Ser	204			B	B				0.57	-0.19			F	1.44	1.00
Glu	205			B	B				0.49	0.00				0.26	0.32
Val	206			B	B				0.79	0.91				-0.32	0.32
Thr	207			B	B				0.90	0.40				-0.60	0.42
Cys	208				B				0.84	0.01				0.10	0.47
Trp	209					T	T		1.14	0.44				0.20	0.63
Gln	210					T	T		0.26	-0.20				0.70	0.75
Arg	211	A				T	T		1.22	0.00			F	0.85	0.98
Gly	212					T	T		1.53	-0.57			F	1.70	1.83
Glu	213		A			T	T		1.61	-1.49			F	1.30	1.75
Ile	214		A			T	T		1.23	-1.39			F	1.15	0.91
Arg	215		A			T	T		0.64	-0.81			F	1.15	0.49
Asp	216		A			T	T		0.53	-0.74			F	1.15	0.29
Ala	217					T	T		0.57	-0.74				1.00	0.71
Cys	218			B					0.36	-0.94				0.80	0.52
Ala	219					T	T		0.93	-0.51				1.00	0.48
Glu	220					T	T		0.61	-0.03			F	0.85	0.89
Thr	221							C	0.40	-0.10			F	1.14	2.00
Pro	222					T			1.03	-0.24			F	1.68	3.06
Thr	223							C	1.49	-0.74			F	2.32	3.53
Pro	224						T	C	2.12	-0.31			F	2.56	3.78
Pro	225					T	T	C	1.31	-0.80			F	3.40	4.89
Lys	226								1.32	-0.54			F	2.86	2.80
Pro	227	A							1.58	-0.64			F	2.32	2.42
Lys	228	A							1.19	-1.07			F	1.78	3.13
Leu	229	A			B				0.51	-0.71			F	1.24	1.36
Ser	230			B	B				-0.09	-0.03			F	0.45	0.62
Lys	231			B	B				-1.02	0.23				-0.30	0.25

Res	Pos	Gami Alpha	Chou- Alpha	Gami Beta	Chou- Beta	Gami Turn	Chou- Turn	Gami Coil	Kyle- Hydro	Modo- Hydro	Eisen Alpha	Eisen Beta	Kami Flex	James Aniq	Finn Sura
Phe	232			B	B				-1.11	0.91				-0.60	0.22
Ile	233			B	B				-1.46	0.61				0.60	0.22
Leu	234			B	B				-1.46	0.61				-0.60	0.14
Ile	235			B	B				-1.74	1.30				0.60	-1.14
Ser	236	A			B				-2.58	1.01				-0.60	0.20
Ser	237	A			B				-2.79	1.01				-0.60	0.17
Leu	238	A			B				-2.71	1.01				-0.60	0.20
Ala	239	A			B				-2.50	1.01				-0.60	0.12
Ile	240	A			B				-2.47	1.24				-0.60	0.09
Leu	241	A			B				-2.47	1.50				-0.60	0.08
Leu	242	A			B				-2.98	1.20				-0.60	0.11
Met	243			B	B				-2.98	1.39				-0.60	0.13
Val	244	A			B				-3.20	1.39				-0.60	0.13
Ser	245	A			B				-3.12	1.39				-0.60	0.13
Leu	246	A			B				-2.61	1.39				-0.60	0.11
Leu	247	A			B				-2.61	1.16				-0.60	0.19
Leu	248	A			B				-2.30	1.20				-0.60	0.12
Leu	249	A			B				-1.40	1.73				-0.60	0.15
Ser	250	A			B				-1.91	1.04				-0.60	0.36
Leu	251	A			B				-1.39	1.04				-0.60	0.36
Trp	252	A			B				-0.47	1.27				-0.60	0.46
Lys	253	A			B				-0.51	0.59				-0.60	0.58
Leu	254	A			B				0.34	0.84				-0.60	0.61
Trp	255	A			B				0.69	0.16				-0.15	1.16
Arg	256	A			B				0.80	-0.76				0.75	1.16
Val	257	A			B				0.28	0.03				0.15	1.21
Lys	258			B	B				-0.66	0.03				-0.30	0.95
Lys	259			B	B				-0.06	-0.20				0.30	0.34
Phe	260			B	B				-0.07	0.23				-0.30	0.71
Leu	261			B	B				-1.03	-0.03				0.30	0.48
Ile	262			B	B				-0.39	0.81				-0.60	0.18
Pro	263			B	B				-0.43	1.04				-0.10	0.32
Ser	264			B	B				-0.69	0.26				0.65	0.54
Val	265							C	0.06	0.00			F	1.30	1.41
Pro	266							C	0.57	-0.69			F	2.50	1.82
Asp	267					T	T	C	0.57	-0.73			F	3.00	1.82
Pro	268					T	T		0.08	-0.43			F	2.60	1.72
Lys	269					T	T		0.17	-0.29			F	2.15	0.98
Ser	270			B	B				0.68	-0.29			F	1.45	0.89
Ile	271			B	B				0.08	0.14				0.20	0.57
Phe	272			B	B				-0.62	0.40				-0.20	0.24
Pro	273			B	B				-0.41	1.19				-0.20	0.15
Gly	274							C	-1.34	0.80				0.00	0.38
Leu	275			B	B				-1.08	0.80				-0.20	0.30
Phe	276								-0.19	0.51				-0.40	0.27
Glu	277	A							0.17	0.49				-0.40	0.47
Ile	278	A							0.38	0.49				-0.40	0.56
His	279	A							0.02	0.20				0.25	1.04
Gln	280							C	0.83	0.20			F	0.45	0.52
Gly	281							C	1.53	0.60			F	0.30	1.29
Asn	282							C	1.24	-0.09			F	1.20	1.84
Phe	283		A					C	1.24	0.33			F	0.05	1.00
Gln	284		A					C	0.97	0.61				-0.16	0.71
Glu	285		A	B	B				0.97	0.67				-0.12	0.63
Trp	286		A	B	B				1.00	0.27				0.57	1.22
Ile	287		A	B	B				1.00	-0.03				1.41	1.02
Thr	288					T			1.70	-0.03			F	2.40	1.02
Asp	289					T			0.84	0.37			F	1.56	1.56
Thr	290	A							0.28	0.10			F	0.92	1.65
Gln	291	A	A						0.51	-0.09			F	1.08	1.15
Asn	292	A	A						0.59	-0.07			F	0.69	0.94
Val	293	A	A						0.87	0.61				0.50	0.54
Ala	294	A	A						0.91	0.63				-0.60	0.42
His	295	A	A						0.62	0.23				0.30	0.52
Leu	296	A	A						0.03	0.44				-0.60	0.70
His	297	A	A						-0.31	0.30				-0.30	0.70
Lys	298	A	A						-0.04	0.23				-0.30	0.51
Met	299	A	A						0.54	0.23				-0.30	0.62
Ala	300	A	A						0.58	-0.46				0.30	0.73
Gly	301	A	A						1.39	-0.56				0.60	0.69
Ala	302	A	A						1.12	-0.56			F	0.90	1.20
Glu	303	A	A						0.73	-0.79			F	1.20	1.60
Gln	304	A	A						1.12	-0.86			F	1.50	1.60
Glu	305							C	1.71	-0.88			F	2.00	2.44
Ser	306							C	2.06	-1.36			F	2.30	2.44
Gly	307					T		C	2.43	-1.36			F	3.00	2.44
Pro	308					T		C	1.62	-1.33			F	2.70	2.18

Res	Pos	Gam- Alpha	Chou- Alpha	Gam- Beta	Chou- Beta	Gam- Tum	Chou- Tum	Gam- Cyl	Kyle- Hydro	Hopp- Hydro	Eisen Alpha	Eisen Beta	Karpi Flex	James Anno	Emre Surfa
Glu	309							C	0.77	-0.64			F	2.40	1.34
Glu	310	A							-0.09	-0.39			F	1.60	1.01
Pro	311	A							0.21	-0.17			F	0.95	0.48
Leu	312		A						0.26	-0.20				0.30	0.48
Val	313	A	A						-0.63	0.49	*			-0.60	0.23
Val	314	A	A						-0.56	0.99	*			-0.60	0.15
Gln	315	A	A						-0.90	0.56	*			-0.60	0.06
Leu	316	A	A						-0.69	0.36	*			-0.30	0.71
Ala	317	A	A						-0.47	0.29	*			0.45	1.65
Lys	318	A	A						0.39	-0.43	*		F	0.45	0.96
Thr	319	A	A						0.94	-0.93	*		F	0.90	2.03
Glu	320	A	A						0.73	-1.13	*		F	0.90	2.69
Ala	321	A	A						1.66	-1.20	*		F	1.24	2.08
Glu	322	A	A						1.64	-1.20	*		F	1.58	2.82
Ser	323	A					T		0.79	-1.07	*		F	2.32	1.61
Pro	324	A					T		1.10	0.39	*		F	2.36	1.32
Arg	325					T	T		0.89	-0.89	*		F	3.40	1.27
Met	326	A					T		1.48	-0.46	*			2.21	1.46
Leu	327	A					T		1.17	-0.44	*		F	1.82	1.64
Asp	328	A					T		1.47	-0.39	*		F	1.68	1.21
Pro	329	A					T		1.68	-0.39	*		F	1.34	2.11
Gln	330	A					T		1.61	-1.00	*		F	1.30	4.44
Thr	331	A					T		2.21	-1.69	*		F	1.30	5.31
Glu	332	A	A						2.43	-1.69	*		F	0.90	5.95
Glu	333	A	A						2.13	-1.61	*		F	0.90	3.47
Lys	334	A	A						2.00	-1.63	*		F	1.15	3.22
Glu	335	A	A						1.66	-1.69	*		F	1.40	1.84
Ala	336	A					T		1.67	-1.26	*		F	2.05	1.05
Ser	337	A					T		0.86	-0.87	*		F	2.15	0.71
Gly	338					T	T		0.86	-0.19	*		F	2.50	0.34
Gly	339					T	T		0.00	0.21	*		F	1.65	0.58
Ser	340							C	-0.21	0.40	*		F	0.70	0.35
Leu	341								0.34	0.44	*			0.30	0.55
Gln	342			B					0.64	0.51	*			-0.15	0.76
Leu	343			B					0.78	0.49	*			-0.40	0.98
Pro	344			B					0.31	0.53	*			-0.25	1.84
His	345			B					0.61	0.53	*			-0.40	0.88
Gln	346			B					1.08	0.53	*		F	0.03	1.84
Pro	347			B					0.73	0.27	*		F	0.46	1.18
Leu	348					T	T		1.54	0.27	*		F	1.04	0.86
Gln	349					T	T		0.90	-0.23	*		F	1.77	0.83
Gly	350					T	T		0.08	0.01	*		F	1.30	0.40
Gly	351			B			T		-0.23	0.23	*		F	0.77	0.36
Asp	352			B	B				-0.91	0.03	*		F	0.24	0.30
Val	353			B	B				-0.44	0.31	*		F	0.11	0.21
Val	354			B	B				-0.79	0.31	*			-0.17	0.21
Thr	355			B	B				-1.14	0.31	*			-0.30	0.13
De	356			B	B				-1.11	1.10	*			-0.60	0.15
Gly	357			B	B				-1.81	0.94	*			-0.60	0.28
Gly	358			B	B				-1.81	1.09	*			-0.60	0.17
Phe	359			B	B				-1.56	1.24	*			-0.60	0.18
Thr	360			B	B				-1.24	1.17	*			-0.60	0.18
Phe	361			B	B				-0.36	1.14	*			-0.60	0.29
Val	362			B	B				0.10	0.71	*			-0.32	0.57
Met	363			B	B				0.14	-0.07	*			0.86	0.77
Asn	364					T	T		0.60	-0.17	*			1.69	1.19
Asp	365					T	T		0.06	-0.20	*			2.37	2.51
Arg	366					T	T		0.17	-0.20	*		F	2.80	1.88
Ser	367	A					T		0.21	-0.31	*			1.97	1.18
Tyr	368	A					T		0.42	-0.03	*			1.54	0.58
Val	369		A	B					0.03	0.40	*			-0.04	0.38
Ala	370		A	B					-0.36	0.83	*			-0.32	0.36
Leu	371		A	B					-0.86	0.87	*			-0.60	0.30
Ter	372		A	B					-0.94	0.54	*			-0.60	0.51

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>4</u> line <u>14</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	
Date of deposit February 25, 1998	Accession Number 209641
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable) The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	
<div>For receiving Office use only</div> <div><input type="checkbox"/> This sheet was received with the international application</div> <div>Authorized officer</div>	<div>For International Bureau use only</div> <div><input type="checkbox"/> This sheet was received by the International Bureau on</div> <div>Authorized officer</div>

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>4</u> line <u>12</u>	
<b>B. IDENTIFICATION OF DEPOSIT</b> Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 10801 University Boulevard Manassas, Virginia 20110-2209 United States of America	
Date of deposit March 23, 1998	Accession Number 209691
<b>C. ADDITIONAL INDICATIONS</b> (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> (if the indications are not for all designated States)	
<b>E. SEPARATE FURNISHING OF INDICATIONS</b> (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g. "Accession Number of Deposit")	

<b>For receiving Office use only</b>	<b>For International Bureau use only</b>
<input type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on
Authorized officer	Authorized officer

*What Is Claimed Is:*

1. An isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:

(a) a polynucleotide fragment of SEQ ID NO:1 or a polynucleotide fragment of the cDNA sequence included in ATCC Deposit No: 209641 or 209691;

(b) a polynucleotide encoding a polypeptide fragment of SEQ ID NO:2 or the cDNA sequence included in ATCC Deposit No: 209641 or 209691;

(c) a polynucleotide encoding a polypeptide domain of SEQ ID NO:2 or the cDNA sequence included in ATCC Deposit No: 209641 or 209691;

(d) a polynucleotide encoding a polypeptide epitope of SEQ ID NO:2 or the cDNA sequence included in ATCC Deposit No: 209641 or 209691;

(e) a polynucleotide encoding a polypeptide of SEQ ID NO:2 or the cDNA sequence included in ATCC Deposit No: 209641 or 209691 having biological activity;

(f) a polynucleotide which is a variant of SEQ ID NO:1;

(g) a polynucleotide which is an allelic variant of SEQ ID NO:1;

(h) a polynucleotide which encodes a species homologue of the SEQ ID NO:2;

(i) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h), wherein said polynucleotide does not hybridize under stringent conditions to a nucleic acid molecule having a nucleotide sequence of only A residues or of only T residues.

2. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding a mature form or a secreted protein.

3. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding the sequence identified as SEQ ID NO:2 or the coding sequence included in ATCC Deposit No: 209641 or 209691.

4. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises the entire nucleotide sequence of SEQ ID NO:1 or the cDNA sequence included in ATCC Deposit No: 209641 or 209691.



5. The isolated nucleic acid molecule of claim 2, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.

6. The isolated nucleic acid molecule of claim 3, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.

7. A recombinant vector comprising the isolated nucleic acid molecule of claim 1.

8. A method of making a recombinant host cell comprising the isolated nucleic acid molecule of claim 1.

9. A recombinant host cell produced by the method of claim 9.

10. The recombinant host cell of claim 9 comprising vector sequences.

11. An isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence selected from the group consisting of:

(a) a polypeptide fragment of SEQ ID NO:2 or the encoded sequence included in ATCC Deposit No: 209641 or 209691;

(b) a polypeptide fragment of SEQ ID NO:2 or the encoded sequence included in ATCC Deposit No: 209641 or 209691 having biological activity;

(c) a polypeptide domain of SEQ ID NO:2 or the encoded sequence included in ATCC Deposit No: 209641 or 209691;

(d) a polypeptide epitope of SEQ ID NO:2 or the encoded sequence included in ATCC Deposit No: 209641 or 209691;

(e) a mature form of a secreted protein;

(f) a full length secreted protein;

(g) a variant of SEQ ID NO:2;

(h) an allelic variant of SEQ ID NO:2; or

(i) a species homologue of the SEQ ID NO:2.

12. The isolated polypeptide of claim 11, wherein the mature form or the full length secreted protein comprises sequential amino acid deletions from either the C' terminus or the N-terminus.

13. An isolated antibody that binds specifically to the isolated polypeptide of claim 11.

14. A recombinant host cell that expresses the isolated polypeptide of claim 11.

15. A method of making an isolated polypeptide comprising:  
(a) culturing the recombinant host cell of claim 14 under conditions such that said polypeptide is expressed; and  
(b) recovering said polypeptide.

16. The polypeptide produced by claim 15.

17. A method for preventing, treating, or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of the polypeptide of claim 11.

18. A method for preventing, treating, or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of the polynucleotide of claim 1.

19. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject related to expression or activity of a secreted protein comprising:

(a) determining the presence or absence of a mutation in the polynucleotide of claim 1;

(b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or absence of said mutation.

20. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject related to expression or activity of a secreted protein comprising:

(a) determining the presence or amount of expression of the polypeptide of claim 11 in a biological sample;

(b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or amount of expression of the polypeptide.

21. A method for identifying binding partner to the polypeptide of claim 11 comprising:

(a) contacting the polypeptide of claim 11 with a binding partner; and

(b) determining whether the binding partner effects an activity of the polypeptide.

22. The gene corresponding to the cDNA sequence of SEQ ID NO:2.

23. A method of identifying an activity in a biological assay, wherein the method comprises:

(a) expressing SEQ ID NO:1 in a cell;

(b) isolating the supernatant;

(c) detecting an activity in a biological assay; and

(d) identifying the protein in the supernatant having the activity.

24. The product produced by the method of claim 23.

1/4

441	ATGATGAAATGATGGAGGCGCTGGCTTCTGCTGCTGCGGAGCTGCGCTGCTTCTGCTGCGA	51
442	M S R L V L L Y S A A V T L L S	18
461	AGCTGATGCGCTTTGGGCGAGGAGAGATAGAGAGAGATACAGATTCAGATTCATCTAT	121
462	L W M A L S L S E A A E S T L L L L Y	36
481	TTCATTTTGAAGAGCTGCACTGATGGAATGCGAGGAAATGCTGAGGAGTAACTGAGGAGTAACTG	181
482	F N L E T V Q V T W H A S E Y S R T L S	54
501	AGTTTTCAGTACAGATTCAGGCTGATGAGGCTATGAGGAGTGCAGAGAAATGATTTT	241
502	T F H Y R F N G E E A Y D D T T N Y L L	72
521	GAGGAAGGTCAGACTTGGGCTGGCTGCTAGAGCGGAGGAGAGAGAGATTCATCTAT	301
522	Q E S H T S G T L L D A S Q W R I L Y	90
541	TCTTCATCAGGAATGGGAGCGAAGCTGCTGAGCGAAGCTGCTGAGTGGTTTATTAT	361
542	P E I E N G T H P V F T A S P W M V Y Y	118
561	CTGAAGCGCAATTTGCTGAAGTACGAGATTTTGTGCAAGAGATTCAGTGCAGGCTG	421
562	I K P Y S P K H Y R F S W H Q D A V T V	136
581	AGCTGTTCTGAGCTTCTAGCGGATGCTGCTGATGAGCTTCACTAGCGGAGCGCTTC	481
582	T C S E L S Y T D L L Y E V Q Y R S P F	154
601	GACAGCGAGTGGCAGTCCAAACAGGAAAATACCTGCAACGTCACCATAGAGGCTTGGAT	541
602	D T E W Q S K Q E N T C N V T I E G L D	172
621	ACCGAGAAGTGCTACTCTTTGCTGGGTGAGGCTCAAGGCTATGCGAGGATGTATATGGCCCA	601
622	A E K C Y S F W V R V K A M E D V Y S P	190
641	GACACATACCGAAGCGACTGCTCAGAGCTGACATGCTGCGAGAGAGCGGAGATTGGGAT	661
642	D T Y P S D W S E V T C W Q R G E I R D	218
661	GGCTGTCCAGAGACACCAACGCTGCCAAGCAAGCTGCTGCAAAATTTATTTTAATTTCC	721
662	A C A E T P T P P K P K L S K F I L I S	236
681	AGCTGGGATGCTTCTGATGCTGCTGCTGCTGCTGCTGCTTATGGAAATATGGAGA	781
682	S L A I L L H V S L L L L S L W K I W R	254
701	GTGAAGAAATTTCTGATTCAGGCTGCGAGAGCGGAAATGATGCTTCCCGGGCTGTTT	841
702	N K Y F L I P S V P D P K S I F P G L F	272
821	GAGATACACCAAGGGAAGCTTCAGGAGTGGATCACAGACACCGAGAAGCTGGGCGAGCTC	901
822	E I H Q G N F Q E W I T D T Q N V A H L	290

FIG. 1A

SUBSTITUTE SHEET (RULE 26)

[illegible]

**SUBSTITUTE SHEET (RULE 26)**

[illegible]

Decoration "Decoration #2": Shade (with solid black) residues that match the consensus named "consensus #1" exactly.

**FIG. 2**  
**SUBSTITUTE SHEET (RULE 26)**

4 / 4

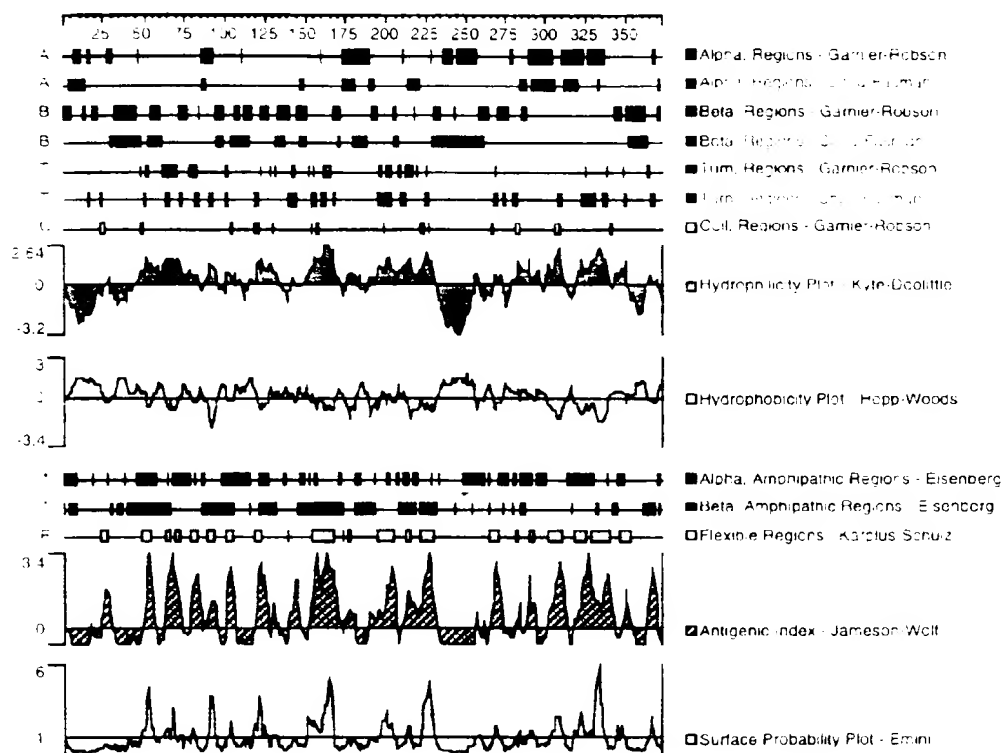


FIG. 3

## SEQUENCE LISTING

(1) Human Protein Kinase C- $\delta$

(11) cDNA sequence (GenBank accession number: U00001)

(12) 10466bp

(13) Unpublished

(14) 1999-03-07

(15) 10466, 505

(16) 1998-05-01

(17) 10466, 505

(18) 1998-03-13

(19) 10

(20) Patent In USA, 200

(21) 1

(22) 1972

(23) DNA

(24) Homo sapiens

(25) 1

(26) CDS

(27) (13-1111135)

(28) 1

gggacgagg gc atg qgg ggg ctg att ctg ctg tgg gga gct gcc gtc ttt 51  
Met Gly Arg Leu Val Leu Leu Trp Gly Ala Ala Val Phe  
1 5 10

ctg ctg gga gcc tgg atg gct ttg qgg caa qga gca gca gaa gga 99  
Leu Leu Gly Gly Trp Met Ala Leu Gly Gln Gly Gly Ala Ala Glu Gly  
15 20 25

gta cag att cag atc atc tac ttc aat tta gaa acc gta cag ctg aca 147  
Val Gln Ile Gln Ile Ile Tyr Phe Asn Leu Glu Thr Val Gln Val Thr  
30 35 40 45

tgg aat gcc agc aaa tac tcc agg acc aac ctg act ttc cag tac aga 195  
Trp Asn Ala Ser Lys Tyr Ser Arg Thr Asn Leu Thr Phe His Tyr Arg  
50 55 60

ttc aac ggt gat gag gcc tat gac cag ttc acc aac tac att ctg cag 243  
Phe Asn Gly Asp Glu Ala Tyr Asp Gln Cys Thr Asn Tyr Leu Leu Gln  
65 70 75

gaa ggt tac act tgg gcc tac ctg cta cag gca gag cag cga gac gac 291  
Glu Gly His Thr Ser Gly Cys Leu Leu Asp Ala Glu Gln Arg Asp Asp  
80 85 90

att ctg tat ttc tcc acc agg aat ggg atg tac gcc gtt ttc acc gca 339  
Ile Leu Tyr Phe Ser Ile Arg Asn Gly Thr His Pro Val Phe Thr Ala  
95 100 105

agt ttc tta atg att tat tac ctg aac ttc att ttc tgg aag tac gta 387



[illegible]

ALL INFORMATION CONTAINED HEREIN IS UNCLASSIFIED  
DATE 08-01-2001 BY 60322 UCBAW

[illegible]

- <210> 2
- <211> 371
- <212> PFT
- <213> Homo sapiens

Met Gly Ala Leu Val Leu Leu Trp Gly Ala Ala Val Phe Leu Leu Gly

Gly Trp Met Ala Leu Gly Gln Gly Gly Ala Ala Glu Gly Val Gln Ile  
20 25 30

Gln Ile Ile Tyr Phe Asn Leu Glu Thr Val Gln Val Thr Trp Asn Ala  
35 40 45

Ser Lys Tyr Ser Arg Thr Asn Leu Thr Phe His Tyr Arg Phe Asn Gly  
50 52 60

Asp Glu Ala Tyr Asp Gln Cys Thr Asn Tyr Leu Leu Gln Glu Gly His  
55 70 75 80

Thr Ser Gly Cys Leu Leu Asp Ala Glu Gln Arg Asp Asp Ile Leu Tyr  
85 90 95

Phe Ser Ile Arg Asn Gly Thr His Pro Val Phe Thr Ala Ser Arg Trp  
100 105 110

Met Val Tyr Tyr Leu Lys Pro Ser Ser Pro Lys His Val Arg Phe Ser  
115 120 125

Tro His Gin Asp Ala Val Thr Val Thr Cys Ser Asp Leu Ser Tyr Gly  
130 135 140

Asp Leu Leu Tyr Glu Val Glu Tyr Arg Ser Pro Phe Asp Thr Glu Trp  
145                      150                      155                      160

Gln Ser Lys Gln Gln Asn Thr Cys Asn Val Thr Ile Glu Gly Leu Asp  
160 170 175

Ala 1 Lys Tyr Tyr Ser Thr Trp Val Arg Val Lys Ala Met Leu Asp  
 1 15 31  
 Thr Gly Gly Ser Trp Tyr Tyr Pro Ser Asp Trp Ser Glu Val Thr Tyr  
 36 52 68  
 Thr Glu Asn Gly Glu Ile Arg Asp Ala Lys Ala Glu Thr Tyr Thr Pro  
 84 100 116  
 Phe Lys Phe Lys Leu Ser Lys Phe Ile Leu Ile Ser Ser Leu Asn Ile  
 122 138 154 170  
 Leu Leu Met Val Ser Leu Leu Leu Leu Ser Leu Trp Lys Leu Trp Arg  
 186 202 218 234  
 Val Lys Lys Phe Leu Ile Pro Ser Val Pro Asp Pro Lys Ser Ile Phe  
 250 266 282  
 Phe Gly Leu Phe Glu Ile His Glu Gly Asn Phe Glu Glu Trp Ile Thr  
 298 314 330  
 Arg Thr Glu Asn Val Ala His Leu His Lys Met Ala Gly Ala Glu Glu  
 346 362 378  
 Glu Ser Gly Pro Glu Glu Pro Leu Val Val Glu Leu Ala Lys Thr Glu  
 394 410 426  
 Ala Glu Ser Pro Arg Met Leu Asp Pro Glu Thr Glu Glu Lys Glu Ala  
 442 458 474  
 Ser Gly Gly Ser Leu Glu Leu Pro His Glu Pro Leu Glu Gly Gly Asp  
 490 506 522  
 Val Val Thr Ile Gly Gly Phe Thr Phe Val Met Asn Asp Arg Ser Tyr  
 538 554 570  
 Val Ala Leu  
 586

<210> 3  
 <211> 363  
 <212> PRT  
 <213> Homo sapiens

<400> 3  
 Met Leu Lys Pro Pro Leu Pro Leu Arg Ser Leu Leu Phe Leu Glu Leu  
 1 5 10 15  
 Asn Glu Asp Ile Gly Gly Lys Pro Gly Thr Gly Gly Asp Phe Phe Leu  
 20 25 30  
 Thr Ser Thr Pro Ala Gly Thr Leu Asp Val Ser Thr Leu Pro Leu Pro  
 35 40 45  
 Lys Val Glu Cys Phe Val Phe Asn Val Glu Tyr Met Asn Cys Thr Trp  
 50 55 60  
 Asn Ser Ser Ser Glu Pro Glu Pro Asn Asn Leu Thr Leu His Tyr Gly

```

65              70              75              80
Tyr Asn Asn Phe Asn Gly Asp Asp Lys Leu Gln Gln Phe Gly His Tyr
      65              70              75              80
Leu Phe Ser Gly Gly Ile Thr Ser Gly Lys Thr Phe Gly Lys Lys Gln
      85              90              95              100
Ile Asn Leu Thr Phe Thr Phe Val Val Gln Leu Gln Asp Phe Asn Gly
      105              110              115              120
His Asn Lys Gln Phe Lys Gln Met Leu Lys Leu Gln Asp Leu Val Ile
      125              130              135              140
Pro Thr Ala Phe Gln Asp Leu Thr Leu Asn Asn Leu Ser Gln Phe Gln
      145              150              155              160
Leu Gln Leu Ser Thr Ser Asn Arg Tyr Leu Asp His Lys Leu Gln His
      165              170              175              180
Leu Val Gln Tyr Arg Ser Asp Arg Asp Arg Ser Thr Thr Gln Gln Ser
      185              190              195              200
Val Asp His Arg His Ser Phe Ser Leu Pro Ser Val Asp Ala Gln Lys
      205              210              215              220
Leu Tyr Thr Phe Arg Val Arg Ser Arg Tyr Asn Pro Leu Cys Gly Ser
      225              230              235              240
Ala Gln His Trp Ser Asp Thr Ser Tyr Phe Ile His Trp Gly Ser Asn
      245              250              255              260
Thr Ser Lys Gln Asn Ile Gln Asn Pro Gln Asn Pro Ser Leu Phe Ala
      265              270              275              280
Leu Gln Ala Val Leu Ile Pro Leu Gly Ser Met Gly Leu Ile Val Ser
      285              290              295              300
Leu Ile Cys Val Tyr Cys Trp Leu Gln Arg Thr Met Pro Arg Ile Pro
      305              310              315              320
Thr Leu Lys Asn Leu Gln Asp Leu Val Thr Gln Tyr Gln Gly Asn Phe
      325              330              335              340
Ser Ala Trp Ser Gly Val Ser Lys Gly Leu Ala Gln Ser Leu Gln Pro
      345              350              355              360
Asp Tyr Ser Gln Arg Leu Cys His Val Ser Gln Ile Pro Pro Lys Gly
      365              370              375              380
Gly Gln Gly Pro Gly Gly Ser Pro Cys Ser Gln His Ser Pro Tyr Trp
      385              390              395              400
Ala Pro Pro Cys Tyr Thr Leu Lys Pro Val Pro
      405              410              415              420

```

&lt;210&gt; 4

&lt;211&gt; 733

&lt;212&gt; DNA

1. *Journal of the American Medical Association*, 1997; 277: 1033-1038.

[illegible]

•211• 5

213 HCM

222

-222- (3)

.223 > Yaad

mp 5er

1 5

..210> 6

..211> 86

212 DNA

-213- Homo sapiens

400 5

tegetteyag atteteeega attettagatt teeteeagatt gattteetee naatgattte 60

Magdalena Alvarado Alvarado

40105 7

00110 17

## 7.10. DNA

[illegible]

[illegible]

Y'taggagat q'taggat q'agagat

400> 15  
 agttaaatat ataaagaaa agtaagagag gta 33

400 16  
gggttagat ctgccatcat ggctttgggg taaggagg 34

400x 17

ggggtttctt gattacaaag ccaggtatga tgggtc

25

<220>  
<221> SITE

<112> 12  
 <113> Xaa equals Ser, Thr, Gl, or Leu

<114>  
 <115> SITE  
 <116>  
 <117> Xaa equals any amino acid

<118>  
 <119> SITE  
 <120> 14  
 <121> Xaa equals Ser or Glu

<122>  
 <123> SITE  
 <124> 15  
 <125> Xaa equals any amino acid

<126> 18  
 Xaa Xaa Trp Xaa Xaa Trp Ser

<127> 19  
 <128> 7  
 <129> PRT  
 <130> Homo sapiens

<220>  
 <221> SITE  
 <222> 12  
 <223> Xaa equals any amino acid

<224>  
 <225> SITE  
 <226> 15  
 <227> Xaa equals any amino acid

<400> 19  
 Thr Xaa Pro Ser Xaa Trp Ser  
 1 5

<210> 20  
 <211> 7  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> 13  
 <223> Xaa equals Pro or Ile

<224>  
 <225> SITE  
 <226> 13  
 <227> Xaa equals any amino acid

<228>  
 <229> SITE



<210> 4  
 <211> Max 40000  
 <212> SITE  
 <213> PRT  
 <214> Homo sapiens

<400> 4  
 Asp Met Val Val Val Val Met Phe  
 1 2 3 4

<210> 5  
 <211> SITE  
 <212> PRT  
 <213> Homo sapiens

<210> 6  
 <211> SITE  
 <212> PRT  
 <213> Max 40000 any amino acid

<400> 6  
 Leu Met Met Val Val Asp Phe  
 1 2 3 4 5 6

<210> 7  
 <211> SITE  
 <212> PRT  
 <213> Homo sapiens

<400> 32  
 Gln Ile Gln Ile Ile Tyr Phe Asn Leu Gln Val Gln Val Thr Trp  
 1 5 10 15

Asn Ala Ser Lys Tyr Ser Arg Thr Asn Leu Thr Phe His Tyr Arg Phe  
 20 25 30

Asn Gly Asp Gln Ala Tyr Asp Gln Cys Thr Asn Tyr Leu Leu Gln Glu  
 35 40 45

Gly His Thr Ser Gly Cys  
 50

<210> 23  
 <211> 30  
 <212> PRT  
 <213> Homo sapiens

<400> 30  
 Arg Arg His Ser Leu Phe Leu Phe Gln Gln Trp Asp Ala Pro Arg Phe  
 1 5 10 15

His Arg Tyr Ser Leu Asp Gly Leu Leu Pro Gln Thr Gln Phe  
 20 25 30

[illegible]

2109 - 25  
 2111 - 181  
 2117 - PBT  
 2118 - Homo sapiens

```

<220>
<221> SITE
<222> (68)
<223> Xaa equals any amino acid

```

```

>220>
>221> SITE
>222> (73)
>223> Xaa equals any amino acid

```

```

-220>
-221> SITE
-222> (88)
-223> Xaa equals any amino acid

```

```

<400> 25
Met Glu Asp Val Tyr Gly Pro Asp Thr Tyr Pro Ser Asp Trp Ser Glu
  1              5              10              15
Val Phe Lys Trp Gln Arg Gly Glu Ile Arg Asp Ala Cys Ala Glu Thr
      20              25              30
Pro Thr Pro Pro Lys Pro Lys Leu Ser Lys Phe Ile Leu Ile Ser Ser
      35              40              45
Leu Ala Ile Leu Leu Met Val Ser Leu Leu Leu Leu Ser Leu Trp Lys
      50              55              60
Leu Trp Arg Xaa Lys Lys Phe Leu Xaa Pro Ser Val Pro Asp Pro Lys
      65              70              75              80

```

1 Ile Pro Pro Gly Leu Phe Met Ile His Cys Gly Arg Phe Gly Val  
 15 20 25 30 35 40 45 50 55 60  
 61 Thr Thr Asp Thr Cys Asn Val Ala His Leu His Tyr Met Ala Gly  
 65 70 75 80 85 90 95 100 105 110  
 121 Thr Cys Cys Ser Gly Ser Thr Thr Phe Leu Val Val Cys Thr Ala  
 125 130 135 140 145 150 155 160 165 170  
 181 Thr Thr Ala Ser Ser Ser Pro Arg Met Leu Asp Pro Cys Thr Cys Gly  
 185 190 195 200 205 210 215 220 225 230  
 241 Thr Ala Ser Gly Gly Ser Leu Cys Leu Pro His Cys Pro Thr Cys  
 245 250 255 260 265 270 275 280 285 290  
 311 Gly Gly Asp Val Val Thr Ile Gly Gly Phe Thr Pro Val Met Asn Asp  
 315 320 325 330 335 340 345 350 355 360  
 381 Arg Ser Tyr Val Ala  
 385

<210> 26  
 <211> 1867  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> (430)  
 <223> n equals a, t, g or c

<220>  
 <221> misc\_feature  
 <222> (416)  
 <223> y equals c or t

<220>  
 <221> misc\_feature  
 <222> (784)  
 <223> m equals a or c

<220>  
 <221> misc\_feature  
 <222> (785)  
 <223> y equals c or t

<400> 26  
 ggcatgggg ggaatgattt tggatgggg aagggcctt ttctactct cagggctgat 60  
 ccttttgggg caaggaagac caggagaag attacagatt caraccatct acttcaattt 120  
 aaaaacqga caggtgcat ggaatggag caatctctt aggaaccaac tgaatttca 180  
 ctacagattc aacggtgata aaggttatga caacttcacc aactaccttc tccagggaag 240  
 tccacttctt ggggtgcttc tagagcaca gaacggagac cacattctct atttctcat 300  
 caggaatgga acgcacccc ttctacagg aagtcctctt atcttttatt acctcaaac 360

caattccctt aacacactta aatttcgttg attttaggaa atggggatgt atttggat 400  
 tttttttt tt tttttt tatgaggttt agttttaggt gttttttt ttgttttt 440  
 tttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt 480  
 ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt 520  
 ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt 560  
 ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt 600  
 ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt 640  
 ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt 680  
 ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt 720  
 ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt 760  
 ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt 800  
 ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt 840  
 ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt 880  
 ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt 920  
 ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt 960  
 ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt 1000  
 ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt 1040  
 ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt 1080  
 ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt 1120  
 ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt 1160  
 ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt 1200  
 ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt 1240  
 ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt 1280  
 ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt 1320  
 ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt 1360  
 ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt 1400  
 ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt 1440  
 ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt 1480  
 ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt 1520  
 ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt ttgttttttt 1560  
 aacttga 1567

&lt;210&gt; 27

&lt;211&gt; 170

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; 891

&lt;223&gt; Xaa equals any amino acid

&lt;230&gt;

&lt;231&gt; SITE

&lt;232&gt; 1132

&lt;233&gt; Xaa equals any amino acid

&lt;240&gt;

<210> 20  
 <211> 20  
 <212> DNA  
 <213> Homo sapiens  
  
 <400> 20  
 cccggttagat ctgccatcat ggggggactg atttata  
  
 <210> 21  
 <211> 31  
 <212> DNA  
 <213> Homo sapiens  
  
 <400> 21  
 ggggggata ggggggata ggggggata g  
  
 <210> 30  
 <211> 4  
 <212> PRT

36

31

12110 Homo sapiens

12111

12112

12113

12114

12115

12116

12117

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US99/05068

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) C07H 21/04; C12N 15/00, 1/20; C07K 1/00; C12P 21/06

US CL. 536/23.5; 435/320.1; 252.3; 69.1; 530/350

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. 536/23.5; 435/320.1; 252.3; 69.1; 530/350

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	TOWNSEND et al. Peptide Mimics of the Cytokine Receptor Common Gamma Chain. Journal of Allergy and Clinical Immunology. 1997, Vol. 99, No. 1 Part 2, page S16, abstract 67.	1-12, 14-16, 21 and 22
A	WATSON, S. et al., editors. The G-Protein Linked Receptor Facts Book. Academic Press. 1994, pages 350 to 355.	1-12, 14-16, 21 and 22

☐ Further documents are listed in the continuation of Box C☐ See patent family annex.

## \* Special categories of cited documents:

\*A\* document defining the general state of the art which is not considered to be of particular relevance

\*B\* earlier document published on or after the international filing date

\*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

\*O\* document referring to an oral disclosure, use, exhibition or other means

\*P\* document published prior to the international filing date but later than the priority date claimed

\*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

\*X\* document of particular relevance, the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

\*Y\* document of particular relevance, the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

\*Z\* document member of the same patent family

Date of the actual completion of the international search

01 JUNE 1999

Date of mailing of the international search report

16 JUL 1999

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

EILEEN B O'HARA

Telephone No. (703) 308-0196

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US99/05068

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2. ☐ Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
  
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
  
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:  
1-12, 14-16, 21 and 22

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest  
☐ No protest accompanied the payment of additional search fees



## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US99/05068

### B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

APS, CAPLUS, MEDLINE, EMBASE, BIOSIS, SWISSPROT, SEQUENCE DATABASES: A\_Geneseq, SPTREMBL, PIR, Issued Patents\_AA, EST, N\_Geneseq, GenEmbl, Issued\_Patents\_NA  
Search terms: receptor, gamma chain

### BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claim(s) 1-12, 14-16, 21 and 22, drawn to nucleic acid molecules, vectors, cells containing recombinant vectors, polypeptides, and a method for making polypeptides.

Group II, claim 13, drawn to an antibody.

Group III, claim 17, drawn to a method for preventing or treating a medical condition by administration of a polypeptide.

Group IV, claim 18, drawn to a method for preventing or treating a medical condition by administration of a polynucleotide.

Group V, claim 19, drawn to a method of diagnosing a pathological condition using genetic analysis.

Group VI, claim 20, drawn to a method of diagnosing a pathological condition by assaying a polypeptide.

Group VII, claims 23 and 24, drawn to a method of identifying activity in a biological assay.

The inventions listed as Groups I to VII do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Pursuant to 37 C.F.R. § 1.475 (d), the ISA/US considers that where multiple products and processes are claimed, the main invention shall consist of the first invention of the category first mentioned in the claims and the first recited invention of each of the other categories related thereto. Accordingly, the main invention (Group I) comprises the first-recited product, polynucleotides encoding cytokine receptor common gamma chain like, a vector, a host cell, a method for producing cytokine receptor common gamma chain like, and cytokine receptor common gamma chain like polypeptide. Further pursuant to 37 C.F.R. § 1.475 (d), the ISA/US considers that any feature which the subsequently recited products and methods share with the main invention does not constitute a special technical feature within the meaning of PCT Rule 13.2 and that each of such products and methods accordingly defines a separate invention.